Direction des Statistiques Démographiques et Sociales

N° F1406

La constitution de l'échantillon démographique permanent de 1968 à 2012

Stéphane Jugnot

Document de travail



Institut National de la Statistique et des Etudes Economiques

INSTITUT NATIONAL DE LA STATISTIQUE ET DES ÉTUDES ÉCONOMIQUES
Série des Documents de Travail
de la
DIRECTION DES STATISTIQUES DÉMOGRAPHIQUES ET SOCIALES

N°F1406

LA CONSTITUTION DE L'ÉCHANTILLON DÉMOGRAPHIQUE PERMANENT DE 1968 A 2012

AUTEUR : STÉPHANE JUGNOT

Document de travail

Septembre 2014

La constitution de l’échantillon démographique permanent de 1968 à 2012

Résumé

L’échantillon démographique permanent (EDP) a été mis en place par l’Insee il y a une cinquantaine d’années pour étudier les comportements démographiques (nuptialité, fécondité, mortalité) selon différentes caractéristiques sociodémographiques, ainsi que la mobilité géographique, professionnelle et sociale. Il est constitué par la juxtaposition d’informations issues de différentes sources statistiques au niveau individuel pour les personnes appartenant à l’échantillon.
La base d’études de 2012, mise à disposition en décembre 2013, comporte 2,7 millions de personnes, pour lesquelles on dispose d’une ou plusieurs observations. Moins de 500 000 d’entre elles ont été recensées en 1968 ; plus de 600 000 en 1999 et moins de 400 000 dans l’enquête annuelle de recensement 2010. Près de 600 000 personnes sont entrées dans l’échantillon à leur naissance.
La façon de définir l’échantillon, les informations prises en compte et la façon de produire la base de données ont évolué au cours du temps avec des implications sur les usages possibles et sur la façon d’utiliser l’EDP. Ce document propose de revenir sur ces différents aspects. Il présente donc les sources utilisées pour alimenter l’échantillon, le champ couvert, l’évolution des informations disponibles. Il revient également sur les chaînes de production de l’EDP et leurs évolutions.
Les utilisateurs trouveront notamment en annexe des statistiques exploratoires, principalement sous forme de graphique.
Ce document est issu d’un travail réalisé de novembre 2011 à aout 2012 sur l’articulation des panels de la direction des statistiques démographiques et sociales de l’Insee. La plupart des données a cependant été actualisée à partir de la base d’étude de l’EDP millésimée 2012, diffusée à la fin de l’année 2013.

Mots-clés : panel, trajectoires, mobilité résidentielle, mobilité professionnelle, fécondité, mortalité, nuptialité.

Abstract
Fifty years ago, INSEE set up the “Permanent Demographic Sample” (« *Echantillon démographique permanent* », EDP) in order to study demographic behaviour (marriage, fecondity, mortality), geographical mobility and occupational mobility according to various sociodemographic factors. The EDP is constructed by grouping together individual information from different data bases (mainly census surveys and civil registration records) for panel members. The selection criterion is the date of birth.
The study database “2012”, made available in December 2013, contains 2.7 million people. However, the coverage of their history is uneven : less than 500,000 people are recorded in the 1968 population census, while more than 600,000 are recorded in the 1999 population census. Nearly 600,000 people entered the sample at birth.
This paper explains how the sample is constructed, what information is collected and how the database is produced. These aspects have evolved over time and need to be fully understood in order to properly use the panel. Exploratory statistics, mainly in the form of graphs, are also provided in the appendix.
This document is the result of work carried out between November 2011 and August 2012 at INSEE. Most of the data presented, however, is based on the study database “2012”.

Keywords : longitudinal study, marriage, fecondity, mortality, geographical mobility, occupational mobility

Table des matières

L'échantillon démographique permanent en bref

Initié à la fin des années 1960, l'échantillon démographique permanent (EDP) permet d'étudier les comportements démographiques (nuptialité, fécondité, mortalité) selon différentes caractéristiques sociodémographiques, ainsi que la mobilité géographique, professionnelle et sociale.

A cet effet, il compile l'ensemble des informations disponibles dans deux sources. D'une part, les bulletins statistiques d'état civil reçus par l'Insee pour les besoins de gestion du répertoire national d'identification des personnes physiques et pour l'établissement des statistiques d'état civil. D'autre part, les informations du recensement de population, puis des enquêtes annuelles de recensement.

Ces informations sont conservées pour les évènements survenus d'abord pour les personnes nées les quatre premiers jours d'octobre. La taille de l'échantillon a été quadruplée à partir de 2004 pour l'état civil et à partir de 2008 pour le recensement. Désormais, les personnes nées le 2, 3, 4 ou 5 janvier ou les quatre premiers jours d'avril ou de juillet, sont également suivies dans l'EDP.

Depuis l'origine, les informations sur l'inscription électorale peuvent également être intégrées dans l'EDP à partir du fichier général des électeurs, également géré par l'Insee. Pendant quarante ans, seuls des rapprochements ponctuels ont été opérés pour réaliser les enquêtes sur la participation électorale. A partir de la base d'études livrée en 2012, les informations sur l'inscription électorale sont systématiquement intégrées.

Le champ couvert est défini par les règles de sélection des évènements. Il correspond approximativement aux personnes résidant en France métropolitaine jusqu'en 2003, puis dans l'ensemble de la France (départements d'outre-mer inclus).

En pratique, **une définition précise du champ ne peut être faite qu'au cas par cas par l'utilisateur, sur la base des critères qu'il retient pour construire la cohorte qu'il souhaite étudier** (personnes nées telles années, personnes recensées telle année…) : le critère géographique est apprécié sur la base du lieu de recensement, donc du lieu de résidence pour le recensement ; pour les évènements d'état civil, il l'est sur la base du lieu de déclaration, donc le plus souvent le lieu d'événement, indépendamment du lieu de résidence (en particulier, celui-ci peut être à l'étranger).

L'EDP permet de réaliser des études de fécondité, de mortalité, de nuptialité différentielle même si **les occurrences des évènements d'état civil sont légèrement sous-estimées puisqu'ils ne sont généralement pas enregistrés lorsqu'ils surviennent à l'étranger.** En particulier, il omet certaines naissances d'enfants et une partie des décès (ce léger biais ne perturbe pas les analyses différentielles et l'étude des évolutions dans le temps).

Toutefois, les utilisateurs doivent être conscients que **l'EDP n'est pas construit pour repérer les personnes mariées à une date donnée** et, qu'**en raison de trous de collecte pour les évènements d'état civil sur la période 1982-1997, les études s'intéressant à la fécondité ou à la nuptialité enregistrée dans l'état civil doivent se limiter aux personnes nées les 1er ou 4 octobre (si les générations étudiées concentrent respectivement beaucoup de naissances ou de mariages sur la période 1982-1997).**

L'EDP permet désormais aussi d'étudier l'inscription électorale. Le projet d'enrichissement en cours de mise en œuvre consistant à intégrer les déclarations annuelles de données sociales, devrait également permettre d'enrichir les analyses sur les trajectoires professionnelles des salariés.

La nature des informations disponibles a évolué au cours du temps, en fonction de l'évolution du contenu des bulletins d'état civil et de ceux du recensement. Pour l'état civil, 1998 et 2008 constituent deux ruptures majeures. Pour le recensement, au-delà du contenu du bulletin lui-même, plusieurs évolutions sont à noter :

1° - De 1968 à 1982, les informations « brutes » correspondent à une saisie des bulletins papier effectuée spécifiquement pour l'EDP ; à partir de 1990, les informations « brutes » correspondent aux informations saisies dans le cadre des chaînes de traitement du recensement.

2°- A partir de 1975, les informations « redressées » dans le cadre des chaînes du traitement du recensement sont intégrées dans l'EDP (y compris les variables imputées pour corriger la non-réponse) ; cette intégration est toutefois **parcellaire pour 1982** (à ce recensement, les ménages comprenant au moins une personne « EDP » n'ont pas été systématiquement intégrés dans l'exploitation « lourde » du recensement, faite sur un échantillon aléatoire du quart des bulletins).

3°- La méthode d'intégration des informations « redressées » permet également d'intégrer des informations issues des autres bulletins de recensement que le bulletin individuel de l'individu « EDP », notamment des informations sur le logement, le ménage, le « conjoint », le « chef de famille », puis le « chef de ménage ». Ces informations complémentaires ont fortement évolué d'un recensement à l'autre, allant vers un enrichissement croissant. A partir des enquêtes annuelles de recensement, les informations sur toutes les personnes des ménages d'individu « EDP » sont conservées.

4°- La mise en place d'une collecte annuelle de recensement sur un échantillon de la population modifie fortement le contexte puisque, désormais, il ne sera plus possible de connaître simultanément les caractéristiques de l'ensemble des personnes « EDP » résidant en France à la date du recensement. Il faudra également recourir aux pondérations du recensement pour assurer la représentativité des cohortes définies par des informations issues des enquêtes annuelles.

Jusqu'au début des années 1990, l'EDP a été géré de façon indépendante, à partir de dossiers papiers actualisés et saisis après chaque nouveau recensement. La gestion de l'EDP a ensuite été intégrée à la sphère de la base des répertoires des personnes physiques (BRPP) que gère l'Insee, si bien que l'intégration au cours du temps de nouvelles informations s'appuie sur un processus d'identification à cette base, à partir des nom, prénom, date et lieu de naissance. Depuis cette intégration, la base d'études est actualisée annuellement, avec des intermittences. En cas de doute sur l'identification, une application dédiée permet de traiter les cas litigieux, ce qui permet d'augmenter le taux d'intégration par rapport à celui qui serait obtenu par un simple appariement automatique.

La rénovation récente de la base des répertoires des personnes physiques a conduit à la suppression de certaines informations non essentielles d'état civil qui étaient auparavant reprises dans l'EDP. Elle a aussi conduit à des pertes d'individus, principalement parmi ceux qui ne sont pas nés en France métropolitaine (40% d'entre eux). Cependant, une partie des informations non « migrées » dans le nouvel environnement a pu être réintégrée après coup dans les bases d'études de l'EDP. En pratique, les personnes « non migrées » sont ainsi réintégrés avec leur trajectoire passée, mais sans possibilité de les suivre ensuite. **Pour certaines utilisations, la nouvelle base d'études ne sera donc représentative que pour le champ des personnes nées en métropole.** Pour l'avenir, toutes les personnes disposant d'un état civil complet dans les bulletins intégrés pourront être intégrées et suivies, quel que soit leur lieu de naissance.

Le projet « EDP+ », amorcé au milieu des années 2000, s'est achevé. Il a principalement consisté, d'une part, à revoir complètement les chaînes de production, les anciennes n'étant pas exploitables dans l'environnement rénové de la BRPP ; d'autre part, à revoir la structure de la base d'études pour substituer un ensemble de tables thématiques à la table unique existant jusqu'alors. Cette restructuration doit faciliter l'ajout ultérieur de nouvelles sources, outre l'intégration systématique d'informations sur l'inscription électorale déjà mise en œuvre dans le cadre du projet « EDP+ ». Le projet « EDP++ », en cours de réalisation, prévoit ainsi l'intégration d'informations sur l'emploi salarié des personnes nées un jour EDP à partir des déclarations annuelles de données sociales (DADS). L'enrichissement par des données fiscales est aussi envisagé, de même qu'une association avec les causes de mortalité du Centre d'épidémiologie sur les causes médicales de décès (CepiDc).

L'EDP est utilisé à des fins d'études et de recherche. Il est mis à disposition des chercheurs après l'avis favorable du comité du secret statistique du Conseil national de l'information statistique (Cnis), celui-ci statuant sur des projets de recherche précis. L'accès s'effectue alors par l'intermédiaire du

Centre d’accès sécurisé distant (CASD[1]), géré par le Groupe des écoles nationales de la statistique (GENES).

L’EDP sert aussi à des productions statistiques régulières, par exemple des statistiques de mortalité par catégorie socioprofessionnelle. Il permet aussi de réaliser des travaux ponctuels d’expertise sur la qualité du recensement et sur les migrations. Les enquêtes sur la participation électorale s’appuient également sur l’EDP.

Un nombre limité de ses informations est associé au panel DADS et, par ce canal, à deux panels de la DREES : l’échantillon inter-régimes des retraités (EIR) et l’échantillon inter-régimes des cotisants (EIC).

[1] Pour en savoir plus, voir le site Internet du CASD (http://www.casd.eu/)

Les sources d'alimentation du panel : les bulletins statistiques de l'état civil, le recensement et le fichier général des électeurs

Historiquement, deux sources alimentent l'échantillon démographique permanent : les bulletins statistiques de l'état civil et les recensements de population. L'ajout d'informations issues du fichier général des électeurs, possible depuis l'origine de l'EDP, n'est que très récent : il a été effectué dans le cadre du projet « EDP+ » et n'intègre donc la base d'études de l'EDP qu'à partir de sa version livrée en 2012.

Bulletins statistiques de l''état civil

Depuis la Révolution, les communes sont responsables de la tenue des registres d'état civil. Conformément à l'instruction générale relative à l'état civil, dont la dernière date du 11 mai 1999, elles sont tenues d'informer l'INSEE de l'enregistrement d'un certain nombre d'actes dans des délais prescrits, au moyen de bulletins statistiques *ad hoc*[2]. Pour l'INSEE, ces bulletins ont généralement une double finalité : d'une part, établir les statistiques d'état civil ; d'autre part, mettre à jour le répertoire national d'identification des personnes physiques (RNIPP), dont l'INSEE assure la gestion depuis 1946.

Les bulletins statistiques, remplis par les officiers d'état civil des communes, sont transmis sur support papier ou électronique en continu et sont traités dans le cadre de la base des répertoires des personnes physiques (BRPP), qui comprend à la fois le RNIPP et le fichier général des électeurs.

Depuis l'initialisation de l'échantillon démographique permanent, les bulletins statistiques de l'état civil ont connu deux révisions importantes, en 1998 et en 2008. Actuellement, neuf types de bulletins statistiques sont utilisés : bulletin de naissance, bulletin de transcription d'un jugement déclaratif de naissance, bulletin de reconnaissance, bulletin d'enfant sans vie, bulletin de mariage, bulletin de décès, bulletin de transcription d'un jugement déclaratif de décès ou d'absence, bulletin de transcription d'un jugement d'adoption plénière, bulletin de mention en marge. Parmi ces bulletins, le bulletin d'enfant sans vie est le seul à n'avoir qu'une finalité statistique. L'annexe 2 les présente de façon détaillée.

Pour les individus et les évènements entrant dans son champ, l'échantillon démographique permanent intègre les bulletins suivants[3] :

- Naissance d'une personne « EDP »,
- Transcription d'un jugement déclaratif de naissance d'une personne « EDP »,
- Transcription d'un jugement d'adoption plénière d'une personne « EDP »[4],
- Mariage d'une personne « EDP »[5],
- Naissance d'un enfant d'une personne « EDP »[6],
- Transcription d'un jugement déclaratif de naissance d'un enfant d'une personne « EDP »[7],
- Transcription de jugement d'adoption plénière par une personne « EDP »[8],
- Naissance d'un enfant sans vie d'une personne « EDP »[9],
- Décès d'une personne « EDP »[10],
- Transcription d'un jugement déclaratif de décès ou d'absence d'une personne « EDP »[11].

---

[2] Ces bulletins sont visés par le CNIS. Le dernier visa couvre la période 2011-2015 (avis de conformité n°36/D131 du 16 avril 2010 ; avis d'opportunité n°14/D130 du 25 janvier 2008).

[3] L'annexe 2 revient plus en détail sur ces bulletins d'état-civil et les situations dans lesquelles ils sont rédigés.

[4] Dans ce cas, les informations initiales sur les parents biologiques ne sont pas conservées.

[5] Quatre mariages au maximum étaient conservés dans l'EDP avant la mise en œuvre de l'EDP+. Il n'y a plus de maximum depuis.

[6] Douze enfants au maximum étaient conservés dans l'EDP avant la mise en œuvre de l'EDP+. Il n'y a plus de maximum depuis. Le nombre maximum porte sur tous les enfants, qu'ils soient repérés par des bulletins de naissance, des bulletins de transcriptions de jugement d'adoption plénière ou des bulletins de transcription d'un jugement déclaratif de naissance.

[7] Voir 4.

[8] Voir 4.

[9] 2 bulletins d'enfant mort-né au maximum étaient conservés dans l'EDP avant la mise en œuvre de l'EDP+. Il n'y a plus de maximum depuis.

[10] Les bulletins de décès n'ont pas été collectés de 1974 à 1989. Au début des années 1990, un rapprochement avec le RNIPP a permis de récupérer une grande partie des décès, mais l'information se limite alors au constat du décès et à la date de décès.

Avant la mise en œuvre du projet EDP+, qui sera présenté plus bas, l'échantillon démographique permanent intégrait également les bulletins suivants :

- Reconnaissance d'une personne « EDP »[12],
- Reconnaissance d'enfants par une personne « EDP »[13],
- Mariage légitimant une personne « EDP »[14],
- Mention en marge de légitimation d'une personne « EDP »[15],
- Autres mentions en marge[16].

Ces bulletins ne sont désormais plus pris en compte dans les nouvelles bases d'études, y compris l'historique disponible : leur intégration aurait nécessité des développements spécifiques coûteux au regard de l'intérêt limité des informations concernées pour de futures études. En particulier, la notion d'enfant légitime n'existe plus dans le Code civil.

Initialement, des bulletins de divorce avaient également été intégrés dans l'échantillon démographique permanent pour les années 1968 à 1974 (ils figurent par exemple dans le panel 1968-1982[17]). Ils ne sont cependant plus présents depuis longtemps dans les fichiers mis à disposition. Il en est de même des bulletins de tentative de conciliation.

Enfin, les enfants sans vie nés un jour « EDP » ne sont pas intégrés dans l'échantillon. Les personnes dont l'existence n'est connue que par un avis de décès ne sont pas non plus intégrées dans les bases de l'EDP mises à disposition des chercheurs.

Recensements de population

Le recensement est réalisé par l'Insee et les communes, avec une implication renforcée de ces dernières depuis la mise en œuvre des enquêtes annuelles de recensement. Je suis surprise de ce « maintenant » . pour moi de tout temps le recensement a été réalisé avec l'aide des communes. Ce qui change maintenant, de mon point de vue, c'est que le partenariat est écrit dans les textes réglementaires.

L'échantillon démographique intègre des informations issues des bulletins des recensements généraux de population de 1968, 1975, 1982, 1990 et 1999 puis, à partir de 2004, des enquêtes annuelles de recensement.

En 1968, seules les informations figurant dans le bulletin individuel des personnes appartenant à l'échantillon démographique permanent ont été intégrées. Depuis le recensement de 1975, des informations figurant dans la feuille de logement et dans les bulletins individuels des autres personnes du ménage sont aussi intégrées.

---

[11] Aucun avant 1993.

[12] Deux reconnaissances au maximum. Ce nombre maximum inclut à la fois les reconnaissances d'une personne « EDP » et les reconnaissances par une personne « EDP ». Ces informations figurent dans les bases d'études antérieures à 2012. A partir de 2008, il n'y a plus de bulletin statistique de reconnaissance.

[13] A partir de 1983. Un saut dans les évènements enregistrés dans l'EDP s'observe en 1998, conduit à s'interroger sur l'exhaustivité pour la période antérieure. Voir également 10.

[14] Un mariage au maximum. Ces informations figurent dans les bases d'études antérieures à 2012, et couvrent les années 1968 à 2005. Cette année là, la notion d'enfant légitime disparait à la suite de l'ordonnance du 4 juillet 2005.

[15] Deux mentions en marge maximum étaient conservées dans l'EDP avant la mise en œuvre de l'EDP+. Il n'y en a plus depuis. Le nombre maximum de mentions pris en compte inclut les « autres » mentions. Des mentions de légitimation par mariage ne sont comptabilisées que sur les périodes 1973 à 1981 et 1990 à 1997.

[16] Deux mentions en marge maximum étaient conservées dans l'EDP avant la mise en œuvre de l'EDP+. Il n'y en a plus depuis. Le nombre maximum de mentions pris en compte inclut les mentions de légitimation par mariage. Dans l'EDP antérieur à la rénovation de la BRPP, les informations associées explicitement aux mentions en marge se limitent à son existence, à sa date et au fait qu'il s'agit d'une légitimation par mariage ou non. La documentation disponible ne permet pas de savoir si seules ces informations étaient intégrées dans l'EDP ou si des informations descriptives renseignées dans la mention étaient aussi utilisées pour renseigner des variables associées à d'autres bulletins dans la base d'études de l'EDP. Dans le cadre de la BRPP rénovée, les mentions modifiant l'état civil de la personne sont prises en compte de facto puisque l'état civil retenu dans la BRPP est celui qui fait référence pour l'EDP.

[17] Voir le manuel d'utilisation de l'échantillon démographique permanent de décembre 1990.

Fichier général des électeurs

Les listes électorales sont gérées par les communes et ne donnent pas lieu à la constitution d'un fichier unique centralisé. Toutefois, l'INSEE gère un fichier général des électeurs[18] destiné notamment à éviter les inscriptions multiples. Il est actualisé à partir des avis d'inscription et des avis de radiation transmis par les communes et des décisions judiciaires.

L'intégration d'informations sur l'inscription électorale est prévue depuis l'origine de l'échantillon démographique permanent. Elle devient effective avec le projet « EDP+ ». Jusqu'alors, des rapprochements ponctuels entre l'EDP et le fichier général des électeurs ont été régulièrement réalisés pour des travaux sur l'inscription et la participation électorale.

Le critère d'échantillonnage : le jour de naissance (de 4 à 16 jours)

L'échantillon démographique permanent ne suit pas une cohorte de personnes initialement identifiées. Il juxtapose les informations des fichiers sources relatives aux personnes nées certains jours de l'année. Un travail spécifique d'identification des personnes est réalisé pour assurer le rattachement des informations aux bons individus au cours du temps (l'absence du NIR dans les fichiers sources de l'EDP, qui oblige à retrouver les personnes à partir de leur état civil déclaré, explique cette différence de traitement entre les deux panels).

Initialement, l'échantillon démographique permanent collectait les informations concernant les personnes nées du 1er au 4 octobre (on parle parfois des « quatre jours EDP »). L'échantillon ainsi constitué regroupait donc *grosso modo* un centième de la population[19]. **Ce choix ne revient pas à un effectuer un tirage strictement aléatoire (certaines études suggèrent par exemple que la période de naissance dans l'année a un effet sur le parcours scolaire, du fait des écarts d'âges à l'entrée en classe préparatoire).**

L'arrêté du 21 décembre 2006 a autorisé le quadruplement de la taille de l'échantillon pour passer à seize jours de naissance. Ils correspondent aux quatre premiers jours de chaque trimestre, à l'exception du mois de janvier, soit les personnes nées du 2 au 5 janvier, du 1er au 4 avril, du 1er au 4 juillet et du 1er au 4 octobre.

Le quadruplement est mis en œuvre à partir de l'année 2004 pour les informations de l'état civil et à partir de 2008 pour les informations issues des enquêtes annuelles de recensement[20].

La base d'études livrée en 2012 est la première à intégrer les douze « nouveaux » jours de naissance.

On parlera par la suite de personnes ou d'individus « EDP » pour désigner les personnes nées un jour retenu dans l'échantillon de l'EDP.

La période couverte : depuis 1967

C'est à l'occasion du recensement de 1968 que l'Insee décide de mettre en place un panel démographique (à la même époque, un panel de salariés voit le jour pour suivre les évolutions salariales ; il deviendra plus tard le panel DADS). L'échantillon démographique permanent rassemble donc les bulletins statistiques de naissances de personnes « EDP » à partir de l'année 1967. Les autres informations d'état civil et les informations issues des recensements sont intégrées à partir de l'année 1968.

---

[18] Loi n°46-1889 du 28 août 1946, repris dans l'article L. 37 du code électoral.

[19] Selon les années, les naissances des quatre premiers jours d'octobre représentent entre 1,00% et 1,25% des naissances enregistrées en France métropolitaine depuis 1968.

[20] L'enquête annuelle de recensement 2008 a été traitée sur les seize jours de naissance mais pour les douze « nouveaux » jours, seule l'identification automatique a été réalisée (les cas non traités automatiquement n'ont pas été examinés par des gestionnaires à l'aide de l'application dédiée).

Pour différentes raisons expliquées dans la partie historique de cette note[21], **il existe** toutefois **des trous de collecte :**

- Les informations issues de l'exploitation « lourde » du recensement de 1982 ne sont disponibles que pour un quart des individus « EDP » recensés en 1982, ceux qui appartenaient à l'échantillon au quart du recensement.

- La plupart des bulletins statistiques de décès couvrant la période 1974 à 1989 n'a pas été intégrée dans l'EDP. La plupart des décès de la période a été repérée ultérieurement par rapprochement avec le répertoire des personnes physiques (seule la date et le lieu de décès sont alors connus).

- Les informations des bulletins statistiques relatifs aux mariages, aux reconnaissances, aux naissances d'enfants de personnes « EDP » (y compris les enfants mort-nés) ne sont pas disponibles pour les personnes nées les 2 ou 3 octobre pour les évènements survenus au cours des années 1982 à 1989[22].

- L'ensemble des bulletins d'état civil de la période 1990-1997 fait défaut pour les personnes nées les 2 ou 3 octobre. Une mise à jour des décès a toutefois été réalisée par rapprochement de l'échantillon démographique permanent avec le répertoire des personnes physiques (seule la date et le lieu de décès sont alors connus).

Le champ couvert : les personnes résidant en France … ou presque

Le champ géographique couvert par l'échantillon démographique permanent est la France métropolitaine jusqu'en 2003. Les départements d'outre-mer sont ensuite inclus.

Le champ géographique est défini à partir d'un critère qui varie selon la source considérée. En particulier, il ne s'agit pas systématiquement du lieu de résidence, même si l'essentiel des évènements intégrés à l'EDP ont eu lieu en France pour des personnes résidant en France :

- **Pour le recensement**, le champ géographique est défini à partir de la commune de recensement, donc la commune de résidence habituelle à la date du recensement.

- **Pour les évènements d'état civil**, trois localisations peuvent être associées : le lieu de l'évènement, le lieu de l'enregistrement de l'acte et le lieu de résidence. Pour les évènements survenus en France, l'enregistrement de l'acte s'effectue normalement dans la commune où l'événement a eu lieu. Ces évènements survenus et enregistrés en France peuvent concerner des personnes résidant à l'étranger. Symétriquement, lorsqu'ils surviennent à l'étranger les naissances d'enfants ayant au moins un parent EDP résidant en France ou le mariage d'une personne EDP résidant en France ne sont généralement pas connus comme événement d'état civil par l'EDP même s'ils peuvent donner lieu à un acte enregistré à l'ambassade[23]. Toutefois, les décès de Français survenus à l'étranger, lorsqu'ils sont enregistrés, intègrent la base des répertoires des personnes physiques, puis l'EDP.

- **Pour le fichier général des électeurs**, seule la commune d'inscription est disponible et fait donc référence. Celle-ci peut différer de l'adresse de résidence puisque toute personne

---

[21] Voir la partie consacrée à « l'histoire de la constitution du panel ».

[22] Seules quelques directions régionales ont intégré ces bulletins.

[23] Selon les cas, les évènements d'état civil concernant des ressortissants français à l'étranger (naissances, mariages, reconnaissances, décès) peuvent être déclarés et enregistrés directement à l'ambassade, l'ambassadeur agissant en tant qu'officier d'état civil. Dans ce cas, ils sont enregistrés par le service central d'état civil du ministère des Affaires étrangères qui centralise depuis 1965 l'ensemble des registres consulaires. Sinon, les évènements d'état civil sont enregistrés selon les procédures du pays concerné. Dans ce cas, les actes d'état civil étrangers peuvent ensuite être transcrits sur les registres d'état civil français. Le service central du ministère des Affaires étrangères n'informe l'Insee que des décès. Cependant, certains actes donnent également lieu à une mention en marge dans la mairie du lieu de naissance si la personne concernée est née en France, mention en marge qui peut alors donner lieu à l'établissement d'un bulletin de mention en marge transmis à l'Insee.

majeure jouissant des droits civiques et répondant à la condition de nationalité[24] peut s'inscrire, au choix, dans sa commune de domicile, sa commune de résidence[25] ou toute commune où elle est assujettie à des contributions communales directes depuis cinq ans au moins. De plus, l'inscription est permanente. En cas de déménagement, la radiation n'est pas automatique tant que l'électeur n'entreprend pas de lui-même les démarches de réinscription dans sa nouvelle commune de résidence. Enfin, les électeurs inscrits au registre des Français établis hors de France peuvent s'inscrire dans une commune française sans y résider[26].

Dans ce contexte, plutôt que de définir le champ de l'EDP de façon générale, il peut être plus pertinent de le définir au cas par cas, étude par étude, en fonction des critères retenus pour définir la cohorte étudiée. Par exemple, « personnes résidant en France à l'enquête annuelle de recensement de telle année », « personnes nées en France métropolitaine de telle année à telle année », « personnes inscrites sur les listes électorales en France métropolitaine telle année », etc.

[24] Le fichier général des électeurs couvre également les ressortissants étrangers de l'Union européenne qui s'inscrivent sur les listes complémentaires pour voter aux élections municipales ou européennes (ils sont alors traités pour l'EDP comme les nationaux, sans distinction particulière). Il couvre aussi les inscriptions des Français de l'étranger sur les listes consulaires (ils ne sont pas intégrés à l'EDP).

[25] Le « domicile » est une notion juridique qui, notamment, n'implique pas une résidence habituelle depuis un temps minimum, à la différence du lieu de « résidence ». Le plus souvent, lieu de résidence et lieu de domicile sont identiques. Exemples de situation où ils peuvent différer : un actif travaillant et résidant la semaine dans une autre commune que le domicile familial ; un étudiant domicilié chez ses parents et résidant dans la commune du lieu d'études.

[26] Les textes listent les communes possibles : leur commune de naissance, leur dernier domicile ou leur dernière résidence ; la commune où est né, est inscrit ou a été inscrit un de leurs ascendants ; la commune où est inscrit ou a été inscrit un de leur parent jusqu'au 4[e] degré ; la commune où est inscrit leur conjoint.

Les informations disponibles : des informations à géométrie variable et à trous

Les manuels d'utilisateur de l'EDP successifs présentent de façon détaillée l'ensemble des variables disponibles dans les différentes bases d'études. Le lecteur pourra donc s'y reporter.

De façon générale, la nature des informations disponibles varie dans le temps, d'une part en fonction de l'évolution des informations contenues dans les bulletins du recensement et dans les bulletins d'état civil ; d'autre part, en fonction des choix de gestion réalisés. L'évolution des nomenclatures complique également l'utilisation longitudinale. Les variables créées spécifiquement pour l'EDP sont peu nombreuses. En particulier, il n'y a pas de variables qui permettraient de sélectionner de façon homogène dans le temps certaines populations selon leurs caractéristiques sociodémographiques, par exemple selon leur origine nationale, leur origine sociale ou leur niveau de diplôme au recensement. En construire ne va cependant pas de soi puisque l'évolution des questions ou des nomenclatures reflète aussi l'évolution des réalités qu'il s'agit d'appréhender, par exemple pour les diplômes ou la catégorie socioprofessionnelle (des tables de correspondance ont cependant été élaborées pour ces deux dernières variables pour assurer une certaine cohérence d'une étude à l'autre).

Par ailleurs, **des trous d'observation peuvent perturber l'appréhension des trajectoires démographiques individuelles que l'EDP permet d'étudier** : des décès, des mariages ou des naissances d'enfants ne sont pas repérés pour les études de mortalité, de nuptialité ou de fécondité différentielle ; la catégorie socioprofessionnelle ou le lieu de résidence ne sont pas toujours connus pour les études de mobilité sociale ou de mobilité résidentielle, etc. Certains de ces trous résultent du champ couvert par les sources alimentant l'EDP (couverture très partielle des évènements d'état civil survenus à l'étranger). D'autres, d'arbitrage sur les moyens disponibles pour intégrer les informations dans l'EDP (interruption de l'intégration des bulletins de décès de 1973 à 1989 et de tous les bulletins d'état civil des personnes nées les 2 et 3 octobre de 1982 à 1997). D'autres enfin, d'évolutions des sources alimentant l'EDP (exploitation « lourde » du recensement de 1982 réalisée sur un sondage au quart ; fin de l'exhaustivité de la collecte du recensement avec la mise en place des enquêtes annuelles de recensement).

- ***A propos des recensements de population***

Pour les recensements de 1968, 1975 et 1982, les bulletins individuels des personnes nées un « jour EDP » ont fait l'objet d'une saisie spécifique pour l'EDP. Les variables concernées sont dites « brutes ». Les informations disponibles concernent les caractéristiques sociodémographiques de la personne (sexe, année de naissance, lieu de naissance, nationalité, lieu de résidence au recensement antérieur, niveau de diplôme, année d'installation en France pour les immigrés), la situation d'activité, le statut, la catégorie socioprofessionnelle, le lieu de travail, le secteur d'activité, le lieu de résidence (commune, département, région).

Pour les recensements de 1975 et 1982, la saisie du numéro du bulletin individuel du recensement des personnes de l'EDP a permis de récupérer les informations issues des travaux d'exploitation du recensement. Les variables concernées sont dites « redressées ». Certaines informations sont disponibles à la fois sous forme « brutes » et « redressées ». Des différences peuvent alors exister, soit du fait d'erreurs de saisie dans l'une ou l'autre des deux opérations (saisie pour l'EDP et saisie pour le recensement), soit du fait des corrections d'anomalies et des imputations de non-réponse effectuées dans le cadre de l'exploitation du recensement, après la saisie des bulletins. Le rapprochement avec le recensement permet surtout de récupérer des informations sur le logement, la composition de la famille et du ménage. Il permet aussi de récupérer une variable de diplôme synthétique, qui isole notamment les « sans diplôme ».

Pour le recensement de 1975, les variables « redressées » sont disponibles pour la quasi-totalité des personnes « EDP » recensées cette année là puisque le recensement a été traité de façon exhaustive. Il l'a été en deux temps : un sondage au cinquième a d'abord été traité pour permettre la publication de premiers résultats nationaux et départementaux ; les quatre cinquièmes suivants ont

été exploités ensuite[27]. Seules **130 communes n'ont pas été traitées par l'exploitation exhaustive** du recensement : 55 communes dans trois départements de Provence-Alpes-Côte d'Azur (dont Nice, Menton, Grasse, Antibes, Arles, Fréjus) ; 70 en Île-de-France (principalement dans les Hauts-de-Seine, la Seine-Saint-Denis, le Val-de-Marne et sept arrondissements de Paris), ainsi qu'à Toulouse, Nantes, Rezé et, de façon partielle, Lyon. Pour ces communes, seule l'exploitation au cinquième a été réalisée. **Les informations issues du rapprochement avec les fichiers du recensement de 1975 sont donc manquantes pour une partie des personnes « EDP » recensées en 1975** (pour 11,3% des cas, ce taux combinant le trou d'exploitation et les échecs d'appariement)[28].

Pour le recensement de 1982, les variables « redressées » ne sont disponibles que pour un quart des personnes « EDP » recensées cette année là : les personnes relevant de l'exploitation « lourde » du recensement (cela concerne toutes les variables de niveau logement, ménage et famille et certaines variables du niveau individuel)[29].

A partir du recensement de 1990, l'ensemble des informations est récupéré directement des chaînes d'exploitation du recensement. **Les variables « brutes » correspondent désormais aux valeurs issues de la saisie réalisée pour le recensement.** Les variables « redressées » restent celles utilisées pour les résultats diffusés du recensement, après détection des anomalies, redressements et imputations de la non-réponse. Les informations intégrées continuent de porter sur les quatre niveaux : individu, ménage, famille, logement.

A partir de 2004, avec la mise en place d'un recensement en continu sur un échantillon rotatif, des informations du recensement sont désormais intégrées chaque année dans l'EDP mais seulement pour un échantillon réduit de la population. L'utilisation de la pondération de l'enquête annuelle de recensement (EAR) est nécessaire pour assurer la représentativité de cohortes définies à partir d'informations issues d'une enquête annuelle donnée. Pour les cohortes définies à partir d'autres critères mais utilisant des informations de plusieurs enquêtes annuelles de recensement, le recours à une pondération spécifique devrait être nécessaire mais son calcul n'est pas forcément évident et devrait s'effectuer à façon.

Quelques variables sont spécifiques à l'EDP :

- Des variables de gestion ;
- Dans certains cas, la codification actualisée de la géographie, en complément de la codification selon le code géographique en vigueur à la date de collecte de la source ;
- Quelques variables construites à des fins d'études dans le passé (repérage des changements de profession et des changements d'établissement employeur entre 1968 et 1975, entre 1975 et 1982 et entre 1982 et 1990 ; repérage des individus du 5e au recensement de 1975).

Les informations du recensement permettent de sélectionner des populations correspondant à certaines caractéristiques à la date d'un recensement donné ou d'observer des évolutions de caractéristiques entre deux recensements. Les informations d'un recensement donné ne permettent pas de caractériser des photographies représentatives à d'autres dates que celles du recensement considéré, d'une part parce que certaines caractéristiques peuvent évoluer dans le temps (statut matrimonial, situation familiale, catégorie socioprofessionnelle ou diplôme atteint, notamment pour les jeunes), d'autre part parce que même pour les caractéristiques stables (origines sociales ou nationales), il faudrait tenir compte des migrations, des naissances et des décès.

---

[27] La variable EB75 permet de distinguer parmi les personnes « EDP » concernées par l'une et l'autre exploitation.

[28] La variable NRP75 permet de repérer les individus pour lesquels le rapprochement a été possible. Elle vaut 1 quand il y a rapprochement et 0 sinon. Elle n'est pas renseignée pour les personnes non recensées (quelques incohérences existent quand on rapproche cette variable du nombre de bulletins du recensement de 1975 associés à la personne - variable NBI75). Une variable de pondération (SOND_75) a été introduite pour assurer la représentativité de l'EDP pour le recensement de 1975. Dans la base exploitée pour la présente note, 455 862 individus ont un poids de 1 ; 11 383 un poids de 5 et 53 un poids de 20 (pour des raisons inconnues). Les individus avec un poids supérieur à 1 représentent donc 2,4% des observations non pondérées et 11,3% des observations pondérées.

[29] La variable ECH82 permet de repérer les individus concernés par l'exploitation au quart (modalité « Q ») des autres (modalité « P ») et la variable NRP82 permet de repérer si la personne a été appariée ou non avec les résultats du recensement. Ces variables ne sont pas renseignées pour les personnes non recensées en 1982. Aucune variable de pondération n'a été introduite.

Figure 1 - Nature des informations disponibles dans l'EDP à partir des recensements généraux de population, puis des enquêtes annuelles de recensement

	1968	1975	1982	1990	1999	EAR
Caractéristiques de l'individu "EDP"						
Saisies spécifiquement pour l'EDP (variables "brutes")	Oui	Oui	Oui	Non	Non	Non
Saisies pour l'exploitation du recensement (variables "brutes")	Non	Non	Non	Oui	Oui	Oui
Corrigées et imputées pour l'exploitation du recensement (variables "redressées")	Non	Oui	Sur 1/4	Oui	Oui	Oui
Caractéristiques du logement de l'individu "EDP"						
Saisies pour l'exploitation du recensement (variables "brutes")	Non	Non	Non	Oui	Oui	Oui
Corrigées et imputées pour l'exploitation du recensement (variables "redressées")	Non	Oui	Sur 1/4	Oui	Oui	Oui
Structure du ménage de l'individu "EDP"						
Corrigée et imputée pour l'exploitation du recensement (variables "redressées")	Non	Oui	Sur 1/4	Oui	Oui	Oui
Caractéristiques des autres personnes du ménage						
Saisies pour l'exploitation du recensement (variables "brutes")	Non	Non	Non	Oui (partiel*)	Oui (partiel*)	Oui (sur toutes)
Corrigées et imputées pour l'exploitation du recensement (variables "redressées")	Non	Oui	Oui (partiel*)	Oui (partiel*)	Oui (partiel*)	Oui (sur toutes)

* Informations sur une partie des personnes du ménage (voir tableau infra).

Figure 2 - Nature des informations disponibles dans l'EDP sur la famille et le ménage où vit l'individu « EDP »

	1968	1975	1982	1990	1999	EAR
Information sur la structure de la famille	Non	Oui	Sur 1/4	Oui	Oui	Oui
Information sur les membres de la "famille" au sens du recensement :						
- le "chef de famille" ("personne de référence de la famille" à partir de 1999)	Non	Oui	Non	Non	Oui	Oui
- le "conjoint du chef de famille" (puis de la "personne de référence de la famille")	Non	Oui	Non	Non	Non	Oui
- le "père de famille"	Non	Non	Sur 1/4	Oui	Oui	Oui
- la "mère de famille"	Non	Non	Sur 1/4	Oui	Oui	Oui
- le "conjoint présumé" de l'individu EDP, s'ils cohabitent	Non	Non	Sur 1/4	Non	Oui	Oui
- les "enfants présumés" de l'individu EDP, s'ils cohabitent	Non	Non	Non	Oui	Oui	Oui
- la "mère présumée" de l'individu EDP, s'ils cohabitent	Non	Non	Non	Oui	Oui	Oui
- le "père présumé" de l'individu EDP, s'ils cohabitent	Non	Non	Non	Non	Oui	Oui
- autres membres de la famille	Non	Non	Non	Non	Non	Oui
Information sur les membres du "ménage"						
- la "personne de référence du ménage"	Non	Non	Non	Non	Oui	Oui
- le conjoint de la "personne de référence du ménage"	Non	Non	Non	Non	Oui	Oui
- autres membres du ménage	Non	Non	Non	Non	Non	Oui

- ***A propos de l'état civil***

Les bulletins d'état civil sont présentés de façon détaillée dans l'annexe 2. Parmi les informations disponibles :

- Les bulletins de naissance permettent de connaître pour les personnes « EDP » nées en France depuis 1967 : l'année de naissance, la commune de naissance et le type de lieu de naissance de la personne (hôpital, domicile,…) ; la profession, la nationalité, l'âge et le lieu de naissance de chacun de ses parents et le fait qu'ils sont mariés ou non ; le lieu de résidence.

- Les bulletins de mariage permettent de connaître le nombre de mariages successifs contractés en France à partir de 1967 par les personnes « EDP », leur date, leur lieu ; la date de naissance, le lieu de naissance et la nationalité de chacun des conjoints, ainsi que leur situation matrimoniale antérieure et leur lieu de résidence (leur profession est aussi disponible jusqu'en 1997). **Les informations sont cependant lacunaires pour les personnes nées les 2 ou 3 octobre en raison du trou de collecte de la période 1982-1997** (du fait de ce trou de collecte, le rang des mariages à nouveau observés à partir de 1998 est aussi biaisé). Par ailleurs, indépendamment du jour de naissance, la remontée des bulletins statistiques de mariage vers l'Insee est celle qui connaît le plus de difficultés, notamment parce qu'alors que les naissances et les décès sont de plus en plus localisés dans un nombre réduit de communes, les mariages restent très diffus sur le territoire, pouvant concerner des petites communes où des oublis de transmission peuvent avoir lieu. Il n'est alors pas facile de les repérer du fait du peu nombre d'actes enregistrés dans ces communes. Le fichier des statistiques d'état civil des mariages est ainsi le seul à faire l'objet d'un redressement en exhaustivité (ces dix dernières années la correction a oscillé entre 3% et 6%). Ce redressement conduit à imputer des mariages dans les fichiers statistiques d'état-civil, qui ne sont naturellement pas intégrés dans l'EDP. Le nombre de mariages y reste donc sous-estimé.

- Les bulletins de naissance et d'enfants sans vie de personnes « EDP » permettent d'approcher le nombre d'enfants des personnes « EDP » et de connaître, leur date de naissance, la commune de naissance, le sexe de l'enfant ; la date et le lieu de naissance, la nationalité et la profession de chacun des parents quand les deux parents sont indiqués dans l'acte ; le fait que les parents sont mariés et dans ce cas, la date du mariage. Les informations sont cependant lacunaires pour les personnes nées les 2 ou 3 octobre en raison du trou de collecte de la période 1982-1997 (du fait de ce trou de collecte, le rang des naissances à nouveau observées à partir de 1998 est aussi biaisé). Ne sont pas non plus pris en compte les bulletins de naissance effectuée sous X.

- Les bulletins de décès permettent de connaître la commune de décès et la date du décès, le lieu de résidence, la situation matrimoniale et la profession à la date du décès. Toutefois, **près de la moitié des décès observés a été repérée à l'issue du rapprochement avec le RNIPP pour corriger les importants trous de collecte pour ces actes (pour tout l'échantillon de 1974 à 1990, puis pour les personnes nées les 2 ou 3 octobre jusqu'en 1997). Dans ce cas, les informations disponibles se limitent essentiellement à la date et au lieu du décès** (i.e. la commune de déclaration du décès). Les décès sont aussi très mal repérés pour les personnes nées avant 1891 (lors du passage du répertoire manuel au répertoire magnétique du RNIPP, seules les personnes nées à partir de 1891 ont été prises en compte).

Dans ce contexte, les études s'intéressant à la fécondité ou à la nuptialité enregistrée dans l'état civil doivent se limiter aux personnes nées les 1er ou 4 octobre si les générations étudiées concentrent respectivement beaucoup de naissances ou de mariages sur la période 1982-1997. L'étude de la mortalité peut utiliser l'ensemble des personnes « EDP ».

Par ailleurs, les informations issues de l'« état civil » dans l'EDP ne permettent pas de définir :

- Quels individus « EDP » sont mariés à une date donnée (les mariages antérieurs à 1968 ou célébrés à l'étranger ne sont pas observés, pas plus que les divorces) ;

- Combien de mariages ont été contractés en France à partir de 1968 par les personnes « EDP » nés les 2 et 3 octobre (pour les personnes en âge de se marier entre 1982 et 1997) ;
- Combien d’enfants nés en France ont eu les personnes « EDP » nés les 2 et 3 octobre (pour les personnes en âge de procréer entre 1982 et 1997).

Figure 3 – Les jours de naissances intégrés dans l’EDP selon l’année et le type de bulletin d’état-civil

<table>
<tr><th rowspan="2">Type de bulletin d’état-civil</th><th>1967</th><th>1968-1973</th><th>1974-1981</th><th>1982-1989</th><th>1990-1997</th><th>1998-2003</th><th>2004-2008</th><th>2008-</th></tr>
<tr><th colspan="6">Parmi les 4 jours d’octobre</th><th colspan="2">Parmi les 16 jours</th></tr>
<tr><td>Naissances</td><td></td><td rowspan="5">Tous</td><td rowspan="4">Tous</td><td rowspan="4">Très lacunaire pour les nés le 2 ou 3 octobre</td><td rowspan="4">Aucun pour les nés le 2 ou 3 octobre</td><td rowspan="5">Tous</td><td rowspan="5">Tous</td><td rowspan="3">Tous</td></tr>
<tr><td>Mariages*</td><td rowspan="4">Aucun</td></tr>
<tr><td>Naissances d’enfants</td></tr>
<tr><td>Reconnaissance</td><td>Sup-primé</td></tr>
<tr><td>Décès</td><td colspan="2">Aucun mais connaissance des décès par la BRPP (bulletins «fantômes»)</td><td>Aucun pour les nés le 2 ou 3 octobre mais bulletins «fan-tômes»</td><td>Tous</td></tr>
</table>

* « Tous » renvoie aux bulletins récupérés par l’Insee. Cependant, certains bulletins de mariage ne sont pas transmis par les mairies à l’Insee.

Comme pour les bulletins du recensement, les informations des bulletins statistiques d’état civil étaient initialement saisies directement pour les besoins propres de l’EDP, indépendamment des traitements réalisés pour les statistiques de l’état civil. L’intégration de la gestion de l’échantillon démographique permanent dans la sphère de la BRPP dans les années 1990 n’a pas modifié cette situation : seules les informations « brutes » étaient intégrées dans l’EDP. Les informations redressées par les chaînes statistiques de l’état civil n’y figuraient pas[30]. La mise en œuvre de la BRPP rénovée modifie le contexte en enchaînant, pour les mutualiser, les contrôles effectués pour les besoins du RNIPP, ceux nécessaires aux statistiques de l’état civil et ceux concernant uniquement l’EDP. Désormais, à partir des évènements survenus dans la deuxième moitié de l’année 2010, la base de production de la base d’études de l’EDP intègre les informations dans leur version « brute » initiale et dans leur version « redressée » pour les statistiques de l’état civil[31]. Cependant, les bases d’études diffusées n’intègrent pas les deux informations (« brutes » et « redressées ») à la différence de ce qui est fait pour le recensement, afin de ne pas trop alourdir les fichiers. Quand le bulletin a été traité pour les besoins des statistiques d’état civil, la priorité est alors donnée aux informations « redressées ». Le choix est réversible puisque les informations « brutes » restent intégrées dans la base de production de l’EDP, en amont de la base d’études.

[30] Ce point doit notamment être pris en compte lorsque l’on rapproche des statistiques calculées à partir de l’EDP avec des statistiques analogues issues des fichiers des statistiques d’état-civil.
[31] Pour les bulletins entrant dans le champ des statistiques de l’état civil, ce qui exclut notamment les évènements intervenus à l’étranger et les bulletins d’enfants sans vie.

Les fichiers mis à disposition : d'une table unique à une série de tables

L'échantillon démographique permanent est constitué d'une base de production évolutive, qui s'enrichit année après année de nouvelles informations. A partir de cette base de production, des bases d'études millésimées sont produites et mises à disposition. Chacune reprend l'ensemble des trajectoires passées des individus « EDP », des modifications pouvant intervenir à la marge quand l'identification d'une personne présente dans les bases déjà produites est modifiée (par exemple, lorsqu'une nouvelle information conduit à corriger sa date de naissance) ou de façon plus substantielle quand certaines informations statistiques sont corrigées, voire (très rarement) retirées, de la base d'études.

De façon générale, une nouvelle base d'études annule et remplace la précédente. Pour disposer des informations les plus complètes, c'est donc la base d'études la plus récente qu'il convient d'utiliser. L'organisation des bases d'études a évolué au cours du temps, passant d'une table unique à une série de tables articulées entre elles par des identifiants.

- ***Jusqu'au millésime « 2006 », livré en 2008 : des tables uniques***

Jusqu'au projet « EDP+ », la base d'études prenait la forme d'une table unique avec une ligne par individu observé.

La base comprend une série de variables fixes, prévues pour accueillir les informations de gestion et, quand elles sont disponibles, les informations des bulletins d'état civil, soit :

- le bulletin de naissance de l'individu « EDP »,
- le bulletin de légitimation de l'individu « EDP » par mariage de ses parents,
- quatre bulletins de mariage,
- douze bulletins de naissance de ses enfants (y compris les transcriptions de jugement et les adoptions plénières),
- deux bulletins de reconnaissance (pouvant être indistinctement la reconnaissance d'un enfant par l'individu « EDP » ou la reconnaissance de l'individu « EDP »),
- le bulletin de décès,
- deux bulletins de mention en marge.

Quand une information est manquante pour l'individu considéré, la variable associée est vide.

Chaque recensement conduit à l'introduction d'une nouvelle série de variables, dont la liste varie d'un recensement à l'autre. Leur nom est suffixé par l'année du recensement.

Initialement, la livraison était ponctuelle, les évènements d'état civil de la période intercensitaire écoulée étant intégrés en même temps qu'un nouveau recensement. La rénovation intervenue au milieu des années quatre-vingt-dix a conduit à une actualisation plus fréquente, l'intégration des évènements d'état civil s'effectuant désormais, plus ou moins régulièrement, chaque année.

Quatorze livraisons sont ainsi disponibles jusqu'en 2008, millésimées par la dernière année d'état civil intégrée : des bases millésimées « 1990 » aux bases millésimées « 2006 ». Il n'y a pas eu de livraison millésimée « 1998 », « 2001 », « 2002 » et « 2004 » et il y a eu deux livraisons pour le millésime « 1999 ».

Jusqu'au millésime « 2003 », chaque livraison comporte trois tables :

- la « base étude » proprement dite ;
- une « base échantillon » correspondant à un échantillon au 78e de la « base étude », plus facile à exploiter avec les technologies de l'époque ;
- une « base statistique » ou « base chemise », réservée au producteur, contenant une centaine de variables de gestion (principales caractéristiques de la personne, compteurs de bulletins, date de création du dossier, indicateurs de qualité, …).

Pour les livraisons millésimées « 2005 » et « 2006 », l'échantillon au 78e a été abandonné.

- ***Une livraison d'attente en 2011, pour le millésime « 2009 » : 20 tables***

En 2011, le millésime « 2009 » de l'EDP a été diffusé dans une architecture transitoire, de façon à proposer aux utilisateurs les informations des enquêtes annuelles de recensement de 2004 à 2009 sans attendre la fin du projet « EDP + ».

La base d'études comprend vingt tables :

- La « table centrale » de 1678 variables a le même dessin que l'ancienne base d'études. Elle est toutefois décomposée en deux tables pour des questions de volume, l'une avec les 500 000 premiers individus, l'autre avec les 521 881 suivants. Elle intègre l'état civil actualisé jusqu'en 2009 et tous les recensements généraux de population.

- Trois tables sont associées à chacune des six premières enquêtes annuelles de recensement, soit dix-huit tables au total. Une première série de tables comporte les informations de niveau « individu » (54 variables, de 260 000 à 280 000 observations selon l'année). Les individus décrits sont les personnes « EDP », mais aussi chacune des personnes résidant dans un ménage qui comporte au moins une personne « EDP ». La seconde série de tables comporte les informations de niveau « famille » (12 variables, de 89 000 à 95 000 observations). La troisième série de tables comporte les informations sur le logement et le ménage (38 variables, de 89 000 à 95 000 observations).

- ***La base d'études « EDP+ »***

Le projet « EDP+ » avait, parmi ses objectifs, une modification complète de la structure de la base d'études pour éviter de poursuivre l'alimentation d'une table unique avec un nombre exponentiel de variables et beaucoup de valeurs manquantes, du fait du passage aux enquêtes annuelles de recensement. C'est pourquoi, la base d'études est désormais éclatée en différents fichiers thématiques. Cette architecture doit aussi faciliter l'ajout de nouvelles sources dans le cadre du projet « EDP++ ». Elle peut aussi faciliter, le cas échéant, des mises à disposition avec des droits d'accès segmentés.

La nouvelle base comprend donc un nombre croissant de tables :

- une table « noyau », avec les principales caractéristiques de l'individu et des informations sur la nature des informations disponibles ;
- une table par type d'évènements d'état civil, soit quatre tables au total : naissance d'une personne « EDP », descendance d'une personne « EDP » (sans nombre maximum), mariage d'une personne « EDP » (sans nombre maximum), décès d'une personne « EDP » ;
- une table sur l'historique de la situation d'inscription électorale ;
- une table par recensement général de population (1968, 1975, 1982, 1990, 1999), soit cinq tables au total ;
- trois tables par enquête annuelle de recensement à partir de 2004 : une table de niveau « logement », une table de niveau « famille » et une table de niveau « individu » (la table « individu » décrit tous les individus des ménages comportant au moins une personne « EDP » ; la table « famille » décrit la famille des personnes « EDP »).

L'annexe 1 décrit de façon plus précise cette nouvelle base d'études.

Au printemps 2013, la première base d'études dans le nouveau format « EDP+ » a été mise à disposition. Cette première base, millésimée « 2011 », comprend 35 tables. La suivante, millésimée « 2012 », mise à disposition en décembre 2013, en comprend 38.

Certaines informations d'état civil disponibles jusqu'alors ne sont plus disponibles dans la nouvelle base d'études :

- les bulletins de mariage légitimant un individu « EDP »,
- les bulletins de reconnaissance d'un individu « EDP »,
- les bulletins de reconnaissance par un individu « EDP »,
- les bulletins de mentions en marge.

L'identifiant utilisé dans la base d'étude de diffusion est non signifiant et attribué aléatoirement[32]. Regénéré chaque année, il ne permet pas de rapprocher les informations contenues dans les bases de deux millésimes différents.

Seuls les évènements associés à une personne « EDP » qui a été identifiée de façon certaine sont intégrés dans la base d'études « EDP+ ». Les évènements se rapportant à des personnes nées un jour « EDP » mais dont l'identité n'a pu être définie de façon certaine, ne sont pas intégrés. La phase d'identification des personnes peut aussi conduire à exclure des personnes si la date de naissance qui figure sur les bulletins en cours d'intégration apparaît erronée et fait l'objet d'une correction qui conduit à exclure la personne des personnes nées un jour « EDP ».

De ce fait, la base d'études de l'EDP ne peut être vue comme une simple extraction des données sources limitées aux personnes supposées « EDP » dans ces sources. Dit autrement, si l'on fait une extraction de tous les individus nés un jour « EDP » dans les fichiers détail du recensement, tous ne se retrouveront pas dans l'EDP. De même, si l'on part des fichiers détail de l'état-civil.

[32] Variable Id_diff (pour l'individu).

Les variables d'identification : l'état civil complet ou un NIR complété

Dans les sources qui alimentent l'échantillon démographique permanent (bulletins d'état civil, bulletins du recensement, avis électoraux), les variables disponibles pour identifier les personnes sont celles qui définissent l'état civil complet de la personne : sexe, nom, prénom, date de naissance (jour, mois, année) et lieu de naissance (commune, pays). La durée de conservation des patronymes (noms et prénoms) dans les sources primaires est toutefois limitée. En particulier, pour les bulletins du recensement, le nom et le prénom sont supprimés au plus tard le 31 décembre de l'année suivant la réception à l'INSEE des fichiers transmis par le prestataire de saisie[33].

Depuis l'origine**, le suivi longitudinal dans l'échantillon démographique permanent s'est donc appuyé sur le NIR, combiné avec le siècle de naissance, du moins pour les personnes pour qui un NIR était disponible.** L'intégration d'une source dans l'EDP impliquait donc au préalable une phase d'identification pour retrouver le NIR de la personne, à chaque fois que cela était possible. Il était ensuite possible de voir si la personne considérée était déjà présente dans l'EDP pour lui associer, dans ce cas, les nouvelles informations disponibles. Le NIR n'était toutefois pas récupérable pour tous puisque seules les personnes nées en France se voient attribuer systématiquement un NIR au répertoire national d'identification des personnes physiques (RNIPP) dès leur naissance. Depuis la mise en œuvre de la BRPP rénovée en 2010, l'usage du NIR n'est plus direct.

Outre les personnes absentes du RNIPP (ou, désormais plus largement, de la BRPP), **des difficultés d'identification peuvent résulter d'informations d'état civil erronées ou incomplètes dans les bulletins d'état civil ou du recensement**. Des erreurs de saisie peuvent aussi intervenir. Pour les femmes mariées, le recours au nom marital est aussi source de difficultés puisque seul le nom de naissance était disponible au RNIPP jusqu'à la mise en œuvre de la BRPP rénovée[34]. L'identification des bulletins suppose donc de traiter des cas litigieux en recourant à des informations complémentaires. Depuis les années 1990, **les litiges sont traités manuellement** par des gestionnaires, à l'aide d'une application dédiée intégrée la sphère « BRPP ».

De façon générale, le suivi est mieux assuré pour les personnes nées en France métropolitaine.

- ***Pour qui le NIR est-il disponible ?***

L'Insee gère le répertoire national d'identification des personnes physiques depuis son origine, aux lendemains de la seconde guerre mondiale. Depuis cette date, toute personne qui nait en France métropolitaine ou dans les départements d'outre-mer y est inscrite et reçoit un numéro d'inscription au répertoire. L'enregistrement systématique des personnes nées sur le territoire s'élargit à Saint-Pierre-et-Miquelon en 1977 puis à l'ensemble de la République française en 2006 (notamment Mayotte, la Polynésie française et la Nouvelle Calédonie)[35]**.** Depuis l'origine, le NIR est également utilisé par les organismes de sécurité sociale. De ce fait, les ayant droit nés à l'étranger ou, jusqu'à récemment,

[33] Article 5 de l'arrêté du 12 janvier 2004 autorisant la mise en œuvre des phases "saisie et exploitation des données collectées" et "contrôle de la cohérence des réponses aux enquêtes" du traitement "Recensement de la population".

[34] Lors de la déclaration auprès de la CNIL du RNIPP au début des années 1980, la possibilité d'intégrer des informations sur la filiation ou le nom marital a été explicitement prévu pour limiter les risques d'homonymie, essentiellement pour les personnes nées hors de France métropolitaine (délibération de la CNIL n° 81-68 du 9 juin 1981 portant avis sur la gestion automatisée d'un répertoire des personnes physiques). La Caisse nationale d'assurance vieillesse disposait de ces informations dans son répertoire. Lors de l'initialisation de la BRPP rénovée en 2011, le rapprochement avec les informations de la Cnav a donc permis de récupérer le nom marital pour une partie des femmes mariées. Il est également désormais collecté par l'intermédiaire des bulletins statistiques de mariage quand l'épouse peut être identifiée à la BRPP. Pour les femmes nées hors de France qui intègrent la BRPP rénovée à l'occasion d'une inscription électorale ou de l'intégration d'une source dans l'EDP, il n'est pas possible de savoir si le nom utilisé est le nom de naissance ou le nom marital.

[35] Décret n°2006-278 du 8 mars 2006 modifiant le décret n° 82-103 du 22 janvier 1982 relatif au répertoire national d'identification des personnes physiques et portant extension à l'outre-mer de l'identification au répertoire.

dans les territoires d'outre-mer recevaient également un numéro d'identification. C'était notamment le cas des travailleurs salariés.

Jusqu'à la mise en œuvre de la BRPP rénovée en 2011

Jusqu'à la mise en œuvre de la BRPP rénovée, les modes de gestion du NIR et de l'EDP ne permettait pas d'utiliser le NIR pour toutes les personnes en disposant. La gestion des dossiers des personnes « EDP » était d'abord essentiellement papier et décentralisée, comme celle du RNIPP, par région selon le lieu de naissance pour la métropole. Trois directions régionales de l'Insee avaient respectivement la responsabilité de la gestion des personnes nées dans les départements d'outre-mer, dans les pays européens et dans le reste du monde. Pour les personnes « EDP » nées en métropole, le NIR était donc utilisé pour identifier les dossiers des personnes « EDP ». **Pour les autres, nées en Outre-mer ou à l'étranger, un pseudo NIR était utilisé dans l'EDP**, construit sur le même modèle que le NIR pour les composantes disponibles (sexe, date de naissance), un code pays et un numéro d'ordre. Le numéro ainsi attribué n'avait donc aucune raison d'être identique à celui attribué le cas échéant pour les organismes de sécurité sociale.

L'informatisation entamée dans les années soixante-dix n'a pas modifié cette dichotomie. Au début des années soixante-dix, la transcription du répertoire a été effectuée sur bande magnétique pour les personnes nées en métropole à partir de 1891, sauf pour celles nées avant 1946 et décédées avant l'informatisation. Ce n'est qu'à partir de1993, avec la mise en place de la première base des répertoires des personnes physiques (BRPP) que les personnes nées en métropole (à partir de 1891), celles nées dans les départements d'outre-mer (à partir de 1900) ou à Saint-Pierre-et-Miquelon (à partir de 1977) ont été gérées dans une base commune, accessible aux gestionnaires de l'EDP. Quant aux personnes nées à l'étranger ou dans les territoires d'outre-mer, l'Insee en a délégué la gestion à la caisse nationale d'assurance vieillesse en 1988 dans le cadre de ce qui est appelé, de façon un peu trompeuse, « la section hors métropole » ou « SHM » [36].

Après la mise en œuvre de la BRPP rénovée

A partir de 2010-2011, la mise en œuvre de BRPP rénovée unifie la gestion de l'ensemble des répertoires. La répartition des rôles entre l'Insee et l'Assurance vieillesse n'est pas remise en cause[37] mais les deux répertoires sont mieux articulés entre eux. **L'EDP peut désormais disposer du NIR pour l'ensemble des personnes qui en disposent, y compris ceux attribués dans le cadre de la SHM.**

La mise en œuvre de la BRPP2 rénovée a eu deux conséquences pour l'EDP :

Premièrement, **à l'initialisation de la BRPP rénovée**, l'ensemble des personnes intégrées dans l'EDP a été ré-identifié dans la nouvelle base des répertoires des personnes physiques. **Seules les personnes identifiées de façon certaine ont été « migrées ».** Le nombre de personnes non migrées est ressorti à 140 000, dont près de 30 000 nées en métropole et 110 000 nées hors métropole. La proportion du nombre de personne nées hors métropole non migrées est ainsi importante (43%). En pratique, la chaîne de production des bases d'étude a donc été adaptée, pour permettre de réintégrer les personnes « non migrées » avec leur trajectoire passée. Pour autant, il ne sera pas possible de les suivre pour les évènements survenus après cette phase d'initialisation. Plus

---

[36] La « section hors métropole », gérée par la Cnav, contient davantage d'informations que le RNIPP pour éviter les risques d'homonymie. Outre le nom patronymique, les prénoms, le sexe, la date et le lieu de naissance, elle peut aussi comprendre le nom usuel ou marital et la filiation (noms et prénoms des père et mère). Dans le cadre de la mise en œuvre de la BRPP rénovée, l'enregistrement du nom marital a été étendu à l'ensemble du RNIPP mais l'enregistrement de la filiation, un temps envisagé, a été finalement exclu. Dans la SHM, pour la construction du NIR, le code du département et celui de la commune sont respectivement remplacés par le code « 99 » suivi d'un code pays sur trois positions. Le code « 99 » était aussi utilisé pour les personnes nées dans les trois départements français d'Algérie avant le 2 juillet 1962. Toutefois, depuis le décret n°2000-910 du 14 septembre 2000, ces personnes peuvent être inscrites avec un code « 91 » à « 94 », selon leur lieu de naissance (sur leur demande dans si elles étaient déjà inscrites au répertoire avant le 29 février 2000, de façon automatique, lorsqu'elles sont inscrites, sinon).

[37] L'Insee récupère toutefois la gestion des personnes nées dans l'ensemble des collectivités d'outre-mer pour couvrir désormais l'ensemble des nés dans un territoire de la République française.

exactement, les personnes réapparaîtront éventuellement en initialisant alors une nouvelle trajectoire. Pour l'utilisateur, tout se passera donc comme s'il y avait deux individus distincts. De ce fait, **pour certaines utilisations particulières s'appuyant sur des cohortes définies par des informations intégrées dans l'EDP avant la mise en œuvre de la BRPP2 et utilisant des informations intégrées après cette date, la nouvelle base d'études ne sera représentative que sur le champ des personnes nées en métropole.**

Deuxièmement, les personnes relevant de la « SHM » ne pourront désormais être intégrées dans l'EDP que si la source à intégrer dispose de l'ensemble des informations d'état civil les concernant (nom, prénom, date et lieu de naissance)[38]. En cas d'informations parcellaires, la personne ne peut pas être intégrée alors qu'elle aurait pu l'être auparavant. Dans l'ancien environnement de la BRPP, l'EDP conservait en effet les patronymes des personnes nées hors de métropole sous une forme simplifiée. Non univoque, elle ne permettait pas de retrouver les noms et prénoms de la personne mais elle donnait une information complémentaire aux informations d'état civil disponibles pour aider à identifier la personne.

- ***Le traitement des cas litigieux***

L'identification des bulletins à intégrer dans l'échantillon démographique permanent comprend deux phases : un traitement automatique puis un traitement manuel des cas litigieux, qui peuvent survenir, y compris pour les personnes nées en France. De façon générale, l'identification consiste à regarder d'abord avec l'état civil complet si la personne existe dans la BRPP[39] ; puis à voir si la personne existe dans l'EDP. Si la personne est retrouvée de manière univoque dans la BRPP et dans l'EDP, l'intégration est automatique. Sinon, le cas est litigieux. Il est alors traité en reprise manuelle.

Il y a litige dans les cas suivants :

- La personne n'est pas retrouvée dans la BRPP. Si elle est née en France, elle devrait y être ; il faut donc déterminer s'il y un trou de collecte dans la BRPP ou un état civil erroné sur le bulletin à intégrer.

- La personne est dans la BRPP mais n'est pas dans l'EDP. Il faut alors déterminer si cette absence est anormale et résulte d'informations erronées sur le bulletin que l'on souhaite intégrer (une personne déclarant être née un jour « EDP » en France métropolitaine après 1967 doit par exemple être dans l'EDP via son bulletin statistique de naissance) ou si cette absence est logique et doit conduire à la création d'un nouvel individu dans l'EDP (parce qu'il s'agit d'un bulletin statistique de naissance un jour « EDP », d'un bulletin de recensement ou d'état civil d'une personne qui pouvait ne pas résider en France auparavant...).

- Les informations d'état civil disponibles donnent trop d'échos dans la BRPP ou dans l'EDP et n'aboutissent pas à un écho unique sans équivoque du fait de l'existence de plusieurs cas voisins. Si le nombre d'échos est limité, il faut alors voir s'il est possible de choisir l'un d'eux. Si le nombre d'échos est très important, il faut alors pouvoir les réduire pour espérer pouvoir traiter le bulletin.

Pour traiter les litiges, les gestionnaires peuvent d'abord utiliser une recherche « standard »[40], qui leur permet d'accéder à l'état civil complet des personnes déjà enregistrées à la BRPP en cas d'échos

[38] Si la personne est déjà inscrite à la SHM, son état civil complet doit permettre de la retrouver. Si la personne n'est pas retrouvée dans la BRPP, dès lors que l'état civil complet est disponible, le traitement de son bulletin pour l'EDP permet d'intégrer cette personne dans la BRPP, puis dans l'EDP. Si une notification de la Cnav arrive postérieurement pour cette personne, les informations d'identification transmises par la Cnav initialement (état civil) se substituent à celles intégrées initialement, voire les complètent (filiation, NIR certifié). En cas d'identification litigieuse, c'est un gestionnaire qui décide si les deux personnes sont les mêmes ou non.

[39] Avant la mise en œuvre de la BRPP rénovée, cette recherche automatique s'effectuait dans le RNIPP pour les personnes nées en Métropole. Pour les autres, la recherche s'effectuait directement dans l'EDP via le sexe, la date de naissance, le lieu de naissance et le nom phonétisé pour les autres. Désormais, tous sont recherchés dans la BRPP.

[40] La recherche « standard » s'appuient sur cinq champs : sexe, nom, prénoms, date et lieu de naissance. Des modifications de ces champs peuvent être testées.

multiples, pour choisir entre ces différents échos. Ils peuvent aussi relâcher ou durcir les critères de façon à augmenter ou diminuer le nombre d'échos dans le cadre de la recherche « étendue », notamment en élargissant la zone géographique de naissance ou la plage des dates de naissances à prendre en compte[41]. Jusqu'à la mise en œuvre de la BRPP rénovée, une recherche "multi critères" était également possible, à partir d'informations disponibles dans l'échantillon démographique permanent mais non constitutives de l'état civil, notamment un indicateur de nationalité, la situation matrimoniale, la catégorie socioprofessionnelle, le dernier domicile connu, le dernier domicile au recensement, l'année et le mois de naissance du dernier enfant. En pratique, ces informations externes liées à des caractéristiques non stables étaient toutefois peu mobilisées.

Pour valider un écho ou choisir entre plusieurs échos proposés, le gestionnaire peut accéder pour chacun d'eux à l'état civil complet disponible dans le répertoire (y compris le nom marital quand il est disponible), à la situation électorale (donc à un lieu de résidence possible), aux informations synthétiques de la « chemise » pour les personnes déjà intégrées dans l'EDP[42] et aux informations contenues dans le bulletin traité. Pour le recensement, ces informations contenues dans le bulletin traité sont toutefois très réduites depuis la mise en œuvre de la BRPP rénovée, puisqu'elles se limitent à l'état civil complet (complété par la commune de recensement). Auparavant, des caractéristiques de la personne et des autres personnes du ménage étaient également disponibles[43].

A l'issue des traitements,

- Les informations qui ont pu être attribuées sans équivoque à une personne déjà présente dans l'« EDP » lui sont associées ;

- Les informations relatives à des personnes nées en France qui ne sont pas retrouvées au RNIPP, ni automatiquement, ni manuellement par les gestionnaires, ne sont intégrées que si, après contact de la mairie de naissance présumée, l'existence de la personne est confirmée, ainsi que sa naissance un jour « EDP » ; la personne est alors préalablement intégrée au RNIPP ;

- Les informations relatives à des personnes nées hors de France qui ne sont pas retrouvées dans la BRPP[44], ni automatiquement, ni manuellement par les gestionnaires, ne sont intégrées dans l'EDP que si l'état civil complet est disponible ;

- Les informations relatives à des personnes aboutissant à plusieurs échos entre lesquels le gestionnaire ne sait pas choisir ne sont pas intégrées dans l'EDP.

- ***L'importance des cas litigieux***

De façon générale, les litiges sont plus fréquents pour les nés hors de France. Ainsi, lors de l'intégration du recensement général de population de 1999, 12 % des bulletins a été traité manuellement : 9 % des bulletins des personnes couvertes par le RNIPP mais 28% des bulletins des autres personnes, les litiges étant par ailleurs plus nombreux pour les femmes que pour les hommes.

---

[41] La recherche « étendue » permet de relâcher :
- le critère de lieu de naissance (rechercher dans tous les arrondissements de la ville, pour Paris, Lyon et Marseille ; dans tous les pays ; dans toutes les communes d'un département donné ; sans indication du lieu de naissance) ;
- le critère de date de naissance (rechercher sur toute une année donnée) ;
- simultanément ces deux critères (rechercher sur toute une année sur tous arrondissements de Paris, Lyon ou Marseille ; rechercher sur toutes les années et tous les départements ou pays - dans ce cas la recherche est étendue par lot de 8 années successives centrées sur l'année indiquée).

[42] La « chemise » correspond à la table Individu dans la version « EDP+ » (voir annexe 1).

[43] L'ajout d'informations sur les autres personnes du ménage, en particulier sur les parents de l'individu « EDP » ou sur ses enfants, en complément de l'état civil complet de l'individu « EDP », supposait la constitution d'un fichier spécifique à partir des fichiers de résultats du recensement. Il était construit par la division des enquêtes et études démographiques. A l'occasion de la mise en œuvre de la BRPP rénovée, la constitution du fichier destiné à l'identification des bulletins du recensement pour l'EDP a été sous-traitée au prestataire de saisie du recensement et le contenu du fichier d'identification a été restreint à l'état civil.

[44] I.e. absentes de l'EDP, du répertoire de la Cnav et du fichier électoral (dans le cadre de la BRPP rénovée).

L'ensemble des bulletins de l'enquête annuelle de recensement 2010 a été traité dans le nouvel environnement de la BRPP. 82 % des bulletins ont pu être traités automatiquement. Réciproquement, 18 % des bulletins ont donc été traités en litiges, soit un peu moins de 75 000 bulletins. 77 % des litiges résultent de problème d'identification (les variables d'état civil sont renseignées et remplies conformément aux normes mais la personne n'est pas retrouvée de façon univoque) ; 19 % des litiges résultent d'un défaut de codification du lieu de naissance (variables non renseignées, incomplètes ou mal codées) et 4 % d'un problème de syntaxe dans le nom ou prénom[45]. Au total, 10 % des litiges n'ont pas été intégrés. 98 % des bulletins récupérés ont donc été intégrés. Les 10 % de litiges non intégrés se décomposent ainsi : 8 % des litiges conduisent à des abandons sans résolution (5 800 bulletins) et 2 % correspondent à des bulletins à ne pas intégrer (1 700 bulletins, dont 1 200 bulletins en double). Même si les personnes nées à l'étranger sont sur représentées dans les abandons, la plupart concerne des personnes nées en France.

Figure 4 – Taux d'insertion automatique des bulletins d'état-civil de 2010

Type de Bulletins	Nombre de bulletins traités	Taux d'insertion automatique	Taux de litiges
Bulletins traités avec les anciennes chaînes de la BRPP			
Naissances	48 000	94	6
Mariages	5 000	86	14
Décès	10 000	98	2
Bulletins traités avec les chaînes rénovées			
Naissances*	59 000	78	22
Mariages	16 000	87	13
Décès*	15 000	82	18
Avis d'inscription électorale	77 000	92	8
Avis de radiation électorale	24 000	91	9
Modifications du répertoire	19 000	69	31
Bulletins d'enfants sans vie	700	25	75
* y compris transcription de jugements et évènements de gestion			

Pour les bulletins d'état civil 2010, ceux arrivés à l'Insee avant la mise en œuvre de la BRPP rénovée en juin 2011 ont été intégrés dans le cadre des anciennes chaînes de traitement. Les autres l'ont été dans les chaînes rénovées. Le tableau ci-dessus (figure 4) présente les taux de litiges pour les différents types de bulletins. Cependant, les données disponibles ne permettent pas de comparaison directe entre les deux chaînes parce que le champ couvert par les indicateurs est plus large pour la partie traitée dans les chaînes rénovées : d'une part, les transcriptions de jugement de naissance ou de décès sont incluses alors qu'elles ne le sont pas pour la partie traitée dans les anciennes chaînes ; d'autre part, le taux de litiges traités dans les nouvelles chaînes porte sur trois niveaux de contrôle successifs (pour la gestion du répertoire, pour les statistiques de l'état civil et pour l'EDP[46]), au lieu d'un seul.

[45] Lors de la saisie des bulletins du recensement, les parties du patronyme peu lisibles étaient remplacées par des astérisques, non lisibles dans la BRPP rénovée. Cette difficulté est maintenant traitée.

[46] Comme indiqué plus haut, la BRPP rénovée mutualise les contrôles davantage que les anciennes chaînes de traitement.

L'articulation avec d'autres sources

Des extraits de l'échantillon démographique permanent ont déjà été associés à d'autres sources externes.

- ***Enquêtes sur la participation électorale***

Les enquêtes sur la participation électorale permettent de produire des résultats nationaux et régionaux sur la récurrence du vote, à partir de la consultation des listes d'émargement des scrutins étudiés pour un échantillon d'électeurs inscrits sur les listes électorales.

Pour cela, l'échantillon démographique permanent a été régulièrement rapproché avec le fichier général des électeurs afin de construire les bases de sondage nécessaires à la réalisation de ces enquêtes (rapprochement automatique, sans recherche ni traitement de cas litigieux). On disposait ainsi des caractéristiques sociodémographiques des électeurs appartenant aux échantillons tirés dans ces bases.

Jusqu'à l'enquête Participation électorale 2008, le rapprochement de l'échantillon démographique permanent avec le fichier général des électeurs s'est effectué sur la base du NIR et du siècle de naissance.

Ces rapprochements sont prévus depuis l'origine de l'EDP même s'ils n'étaient pas systématiques. Avec la mise en œuvre du projet « EDP+ », l'information sur l'inscription électorale est désormais partie intégrante de l'EDP, ce qui devrait faciliter la production des enquêtes sur la participation électorale ainsi que la production régulière de statistiques sur l'inscription électorale en fonction de différentes caractéristiques sociodémographiques.

- ***Panel DADS***

L'arrêté du 24 juin 1998 relatif aux traitements réalisés sur les DADS permet l'enrichissement du panel "DADS" par un nombre restreint d'informations explicitement listées : le diplôme obtenu, les dates de mariage, le nombre et les dates de maternité. Cette autorisation a été élargie à l'ensemble des DADS « grand format » en 2006[47]. Par ce canal, des informations issues de l'EDP sont également intégrées, conjointement à celles du panel DADS, dans deux échantillons de la DREES : l'échantillon inter régimes des retraités (EIR) et l'échantillon inter régimes de cotisants (EIC).

Initié au début des années soixante-dix, le panel « DADS » permet de suivre les trajectoires salariales d'un échantillon d'individus à partir des informations contenues dans les déclarations annuelles des salaires (DAS) puis les déclarations annuelles de données sociales (DADS). Cette déclaration administrative doit être remplie pour tous les salaires versés par l'établissement employeur, indépendamment du lieu de résidence des salariés.

Son champ, la nature des informations disponibles et leur qualité ont évolué au cours du temps en fonction de l'évolution des règles administratives et des chaînes d'exploitation de l'Insee. De ce point de vue, les principales ruptures ont lieu à l'occasion des déclarations relatives à 1976 (« validité 1976 »), à 1993 et 1994 (« validités » 1993 et 1994) et 2002 (« validité 2002 »).

Comme pour l'échantillon démographique permanent, la date de naissance sert de critère d'échantillonnage : les personnes nées en octobre des années paires pour les déclarations relatives aux années 1967 à 2001 (échantillon au 25e), toutes les personnes nées en octobre à partir de 2002 (échantillon au 12e).

---

[47] Le champ de la source DADS couvre les salariés du secteur privé, la fonction publique territoriale et hospitalière. Les DADS « Grand Format » intègre également la fonction publique d'Etat et les salariés des particuliers employeurs.

Les obligations déclaratives couvertes par la DADS[48]

Déclaration	Depuis le	Remarques
1) la déclaration de traitements, émoluments, salaires ou autres rétributions imposables prescrites à l'article 87 du code général des impôts ; 2) la déclaration des commissions, courtages, ristournes, honoraires, droits d'auteurs, rémunérations d'associés et parts de bénéfices et autres versements qui est prescrite aux articles 240 et 241 du même code ; 3) la déclaration des rémunérations versées aux salariés prescrite à l'article R. 243-14 du code de la sécurité sociale ;		Obligations initiales de la **DAS**, issue de la fusion des déclarations fiscales déjà traitées par l'INSEE (points 1 et 2) et d'une déclaration « sociale » (point 3)
4) l'attestation d'activité salariée prévue pour l'application de l'article R. 313-2 du code de la sécurité sociale ;	1985 (?)	La DAS devient **DADS** L'AAS permet notamment de vérifier des seuils de nombre de jours et de nombre d'heures travaillés.
5) la déclaration prévue par l'article 6 bis du décret n° 70-1277 du 23 décembre 1970 susvisé. 6) la déclaration prévue par l'article 3 du décret n° 47-1846 du 19 septembre 1947 ; 7) la déclaration prévue par l'article 4 du décret n° 50-783 du 24 juin 1950 ; 8) la déclaration prévue par l'article 3 du décret n° 85-885 et du décret n° 85-886 du 12 août 1985.	6 janvier 1989	Intégration de la fonction publique territoriale
9) la déclaration relative à l'emploi obligatoire des travailleurs handicapés, des mutilés de guerre et assimilés prévue par l'article L. 323-8-5 du code du travail pour les renseignements énoncés au 1° de l'article R. 323-9 du même code ;	16 mai 1996	Intégration de l'ESE (codification PCS-ESE sur les établissements de plus de 20 salariés)
10) la déclaration prud'homale mentionnée au I de l'article L. 513-3 du code du travail ; 11) la déclaration mentionnée à l'article 15 du décret n° 2004-569 du 18 juin 2004 susvisé [sur la retraite additionnelle de la fonction publique] ;	9 sept. 1999	Intégration du collège électoral et de la section prud'homale.
12) la déclaration adressée à la Caisse nationale des barreaux français par les employeurs d'avocats salariés ;	30 déc. 2006	
13) la déclaration mentionnée à l'article L. 1221-18 du code du travail [pré retraites] ; 14) la déclaration prévue par l'article 8 du décret n° 2007-1796 du 19 décembre 2007 relatif à la cotisation et à la contribution dues pour la couverture des charges de pensions et allocations temporaires d'invalidité des fonctionnaires de l'État, des magistrats et des militaires détachés ainsi que des agents des offices ou établissements de l'État dotés de l'autonomie financière ;	19 avril 2008	
15) la déclaration annuelle des salaires adressée par les employeurs de personnel navigant professionnel de l'aéronautique civile à la caisse mentionnée à l'article L. 426-5 du code de l'aviation civile ; 16)la déclaration mentionnée à l'article 3 de la loi n° 2009-594 du 27 mai 2009 pour le développement économique des outre-mer.	9 mai 2009	

[48] D'après l'article 2 du décret n°85-1343.

Le NIR est la variable d'identification utilisée pour assurer le suivi dans le temps. Il est naturellement exclu des fichiers de diffusion. C'est cette information, complétée par le siècle de naissance, qui permet de rapprocher le panel DADS avec l'échantillon démographique permanent pour les personnes entrant dans le champ du RNIPP, donc les personnes nées en France métropolitaine. Pour les autres personnes, les DADS disposent du NIR attribué à la SHM alors que l'EDP disposait d'un pseudo-NIR différent jusqu'à la mise en œuvre de la BRPP rénovée. Pour ces personnes, le rapprochement des deux sources n'est pas possible. La rénovation de la BRPP modifie le contexte et permet d'envisager un rapprochement pour l'ensemble du champ couvert par les DADS. La représentativité de l'appariement DADS-EDP n'est donc pas assurée pour les personnes nées hors de France métropolitaine quand les cohortes sont définies par des informations intégrées dans l'échantillon démographique permanent avant la mise en œuvre de BRPP2.

Par ailleurs, le NIR n'est pas disponible pour certaines observations des DADS, en amont de la constitution du panel. Afin de pouvoir utiliser celui-ci pour produire des statistiques sur les salaires et leurs évolutions, la source cherche cependant à être représentative de la source « DADS » pour chacune des années. Chaque année, des observations sont donc tirées au hasard parmi celles dont la date de naissance n'est pas connue pour compléter l'échantillon qui alimente ensuite le panel DADS. Ces personnes complémentaires reçoivent un NIR « fictif » qui ne permet pas de les suivre dans le temps. Pour ces personnes, une association avec l'échantillon démographique permanent n'est pas possible.

À côté du panel DADS, un panel « Tous salariés » (ou « grand format ») a été mis à disposition pour la première fois en mars 2012 sur la période 1988-2009. Il intègre les informations des agents de l'État à partir de 1988 ainsi que les particuliers employeurs pour 2009.

- ***Deux exemples de projets ponctuels sur les causes de mortalité***

COSMOP (causes de mortalité, InVS)

Le projet « COSMOP » (Cohorte pour la surveillance de la mortalité par profession) avait pour objectif la mise en place d'un dispositif d'observation régulier sur la mortalité par cause en fonction de l'activité professionnelle. Il s'agissait de suivre la mortalité par profession et secteur d'activité dans le cadre de la surveillance épidémiologique des risques professionnels.

En 2002, l'Institut de veille sanitaire (InVS) a lancé une étude de faisabilité basée sur un enrichissement de l'EDP par les causes de mortalité collectées par le CépiDc de l'Inserm (l'InVS a accès aux causes de mortalité en vertu de la loi 98-535 du 1er juillet 1998 relative au renforcement de la veille[49]). Cette étude porte sur les individus de l'EDP nés en 1974 ou avant. Les informations transmises par l'Insee portent sur l'état vital, connu jusqu'en 1997, l'activité professionnelle aux recensements de 1968, 1975, 1982 et 1990, ainsi que les variables nécessaires à la récupération des causes de décès par le CépiDc (sexe, date et lieu de naissance, date et lieu de décès). Les résultats de cette étude de faisabilité ont fait l'objet d'un rapport détaillé publié par l'InVS en septembre 2006[50].

[49] Cette loi crée l'article L792-2 du Code de la Santé publique qui autorise l'Institut de veille sanitaire à collecter auprès de diverses administrations les données nécessaires à l'accomplissement de ses missions, y compris celles couvertes par le secret médical. Après l'abrogation de cet article lors de la refonte du code de la Santé publique, les conditions de communication d'informations couvertes par le secret médical à l'InVS sont désormais précisées dans la sous-section 5 de la section 1 (« Institut de veille sanitaire ») du chapitre III (« Sécurité, veille et alertes sanitaires »), du titre Ier (« institutions ») du livre IV (« administration générale de la santé »), dans la première partie de la partie réglementaire, soit les articles R1413-21 et suivants.

[50] *Analyse de la mortalité et des causes de décès par secteur d'activité de 1968 à 1999 à partir de l'échantillon démographique permanent,* rapport de l'InvS. Septembre 2006. 159 p. Disponible à la date de rédaction de ce document au lien suivant : http://www.invs.sante.fr/publications/2006/rapport_cosmop/rapport_cosmop.pdf

Projet EDISC (causes de mortalité, Inserm)

Le projet « EDISC » (Évolution des inégalités sociales par causes médicales de décès) a été lancé au début des années 2000, avec des objectifs proches de ceux de Cosmop par l'unité 687 de l'Inserm[51].

Après l'autorisation de la CNIL obtenue par l'Inserm en janvier 2003, l'enrichissement par les causes de mortalité s'est effectué selon des principes analogues à ceux du projet COSMOP. L'enrichissement par les causes de décès a été poursuivi et couvre actuellement les décès jusqu'à 2007 inclus. L'utilisation de cette base doit permettre à l'Inserm de continuer de documenter les inégalités sociales d'incidence des cancers dans le cadre du plan Cancer 2009-2013.

[51] L'unité 687 **«** Epidémiologie des déterminants sociaux et professionnels de la santé », a intégré en 2010 le centre de recherche « Epidémiologie et santé des populations » (URMS 1018).

L'histoire de la constitution du panel EDP : d'une gestion manuelle à une gestion informatisée dédiée intégrée à la sphère « BRPP »

- ***1968-1989 : l'EDP « ancienne manière »***[52]

En 1967, l'Insee décide de mettre en place un « panel démographique » à l'occasion du recensement de 1968, pour un échantillon au centième de la population, tiré à partir de la date de naissance (personnes nées les quatre premiers jours d'octobre). Il s'agissait notamment de construire un outil plus efficace que les enquêtes familles réalisées en 1957 et 1962 pour étudier les comportements de fécondité, pour éviter les biais des enquêtes rétrospectives et élargir le champ couvert au-delà des seules femmes mariées. Il faudra attendre les années quatre-vingts pour voir les premières études publiées, d'abord sur la participation électorale puis sur la mobilité géographique. Ce n'est que dans la deuxième moitié des années quatre-vingt-dix que l'échantillon démographique permanent devient une source plus régulièrement exploitée par les chercheurs.

De 1968 à 1989, l'alimentation de l'échantillon démographique permanent s'appuie sur une gestion manuelle des bulletins. Les bulletins d'état civil et les bulletins individuels de recensement des personnes nées un jour « EDP » sont rassemblés dans des dossiers individuels conservés par la direction régionale de l'Insee de la région de naissance de la personne. Trois directions régionales sont respectivement en charge des dossiers des personnes nées dans les départements d'outre-mer, dans les pays européens et dans les autres pays. Il faut donc que les bulletins circulent entre la direction régionale de traitement (celle du lieu de résidence au recensement par exemple) et la direction régionale de gestion du dossier. Des duplications sont réalisées quand un bulletin concerne plusieurs personnes « EDP » (un enfant né un jour « EDP » ayant un parent né lui aussi un jour « EDP », un mariage entre deux personnes nées un jour « EDP », …). Pour les personnes nées en métropole, l'utilisation du NIR, numéro individuel au répertoire national d'identification des personnes physiques (RNIPP), permet d'identifier les dossiers de manière unique, le NIR étant lui-même retrouvé à partir de l'état civil figurant sur les bulletins (nom, prénom, date et lieu de naissance). Pour les personnes hors de métropole, un NIR fictif était attribué, construit de façon analogue au NIR.

Au départ, l'alimentation des dossiers ne s'effectue pas en continu. Entre deux recensements, les bulletins d'état civil sont stockés dans la direction régionale qui les collecte pour les besoins de gestion du répertoire des personnes physiques et des statistiques d'état civil. Dès qu'un nouveau recensement est réalisé, les bulletins d'état civil et les bulletins individuels du recensement sont transmis aux différentes directions régionales compétentes afin d'être intégrés dans les « chemises » des personnes concernées. Des nouvelles chemises doivent être créées pour les personnes qui ne disposent pas déjà d'un dossier. Les informations sont ensuite chiffrées et saisies pour obtenir un fichier magnétique contenant l'ensemble des informations. La saisie des bulletins faites pour le panel démographique ne profite donc pas des traitements réalisées pour les statistiques d'état civil, d'une part, pour le recensement d'autre part.

Toutefois, à l'occasion du recensement de 1975, l'identifiant du bulletin de recensement est également saisi pour permettre un rapprochement avec les informations codées par les chaînes de traitement du recensement. Les informations brutes, saisies spécifiquement pour l'échantillon démographique permanent, avec les mêmes consignes que pour le recensement, sont ainsi complétées par les informations codées dans le cadre des exploitations du recensement, lesquelles incluent également des redressements et des imputations de données manquantes. Ce rapprochement permet aussi d'introduire des informations des niveaux famille, ménage et logement[53].

---

[52] L'expression est de Dominique Rouault dans son article « L'échantillon démographique permanent a pris un coup de jeune : gestion informatisée et nouvelle base d'études », *Le courrier des statistiques*, n°73, Insee, mars 1995. La plupart des aspects historiques présentés dans cette partie sur la période 1967-1995 est issue de cet article.

[53] Le rapprochement a été réalisé à partir d'un extrait du fichier du recensement de 1975 limité aux personnes nées les quatre premiers jours d'octobre. Il s'est appuyé sur l'identification complète du logement et l'état civil partiel de la personne (jour, année, département ou pays de naissance). Les critères ont ensuite été progressivement relâchés. Le taux de succès atteint près de 90 % (par rapport aux personnes de l'échantillon ayant un bulletin du recensement inséré dans leur dossier EDP). La variable NRP75 permet de repérer les personnes pour lesquels l'appariement a été effectué.

La base constituée est utilisée pour réaliser des travaux internes de nature méthodologique sur la qualité du recensement[54] mais peu d'études externes. On trouve ainsi mention de deux études de Jean Morin utilisant les fichiers produits après le recensement de 1975 : une étude statistique sur l'inscription sur les listes électorales, la première du genre (« Un français sur dix ne s'inscrit pas sur les listes électorale », *Economie et Statistiques*, n°152, Insee, février 1983) et les résultats de la première enquête sur la participation électorale, alors expérimentale (« La participation électorale dans les grandes villes aux élections municipales de 1983 », *Economie et Statistiques*, n°165, Insee, avril 1984).

Après le recensement de 1982, l'exploitation complète du recensement (dite aussi exploitation « lourde ») est restreinte à un sondage au quart de la population. Son rapprochement avec l'échantillon démographique permanent s'en trouve restreint d'autant. Comme pour le recensement de 1975, le rapprochement a été réalisé à partir d'un extrait du fichier du recensement de 1982 limité aux personnes nées les quatre premiers jours d'octobre. Il s'est appuyé sur l'identification complète du logement et l'état civil partiel de la personne (jour, année, département ou pays de naissance). Les critères ont ensuite été progressivement relâchés[55]. Les informations des niveaux familles, ménages et logements ne sont donc disponibles que pour un quart des personnes présentes dans l'EDP et recensées en 1982. Il en est de même des codifications « recensement » au niveau individu, notamment la variable synthétique sur les diplômes. Pour les autres, seules les informations du bulletin individuel saisies pour l'EDP sont disponibles. Le fichier magnétique est livré en 1988, avec les informations des recensements de 1968 à 1982 et l'état civil de 1968 à 1981[56]. Quelques publications et communications dans des colloques utilisent cette base d'études, notamment sur la thématique des mobilités géographiques[57].

- ***1990-1995 : l'informatisation de la gestion***

Au début des années quatre-vingt dix, une rénovation importante du mode de gestion de l'échantillon démographique permanent est entreprise. Il s'agit alors de réduire les manipulations manuelles (tri, échanges, classement des bulletins, saisie), source d'erreurs et d'omission ; de lisser les travaux de mise à jour pour réduire les à-coups de charge de travail générés jusqu'alors par les mises à jours post-censitaires ; d'accélérer la mise à disposition des fichiers actualisés.

Pour cela, plusieurs innovations importantes sont introduites pour réduire au minimum les travaux réalisés spécifiquement pour l'EDP en l'adossant davantage aux chaînes statistiques des sources qui l'alimentent :

- *Mise en place d'une base de données informatiques mise à jour informatiquement en remplacement des dossiers papiers*. En pratique, la gestion de l'EDP intègre la sphère des répertoires des personnes physiques (BRPP), aux côtés du RNIPP et du fichier général des

[54] Par exemple, sur la qualité de la saisie grâce à la double saisie du contenu des bulletins, effectuée d'une part pour le recensement, d'autre part spécifiquement pour l'EDP, ou sur les doubles comptes.
[55] Le taux d'appariement avec les personnes ayant un bulletin du recensement de 1982 concerné par l'exploitation au quart dans leur dossier EDP a été de 87,5% avec l'ensemble des critères d'indentification et atteint 99,7% en relâchant progressivement les critères.
[56] La base d'études est déjà une table SAS composée d'une seule table, comprenant 1 295 variables et 711 038 observations (595 010 personnes identifiées nées en France métropolitaine, 101 917 personnes nées dans les DOM ou à l'étranger et 14 111 personnes nées en France métropolitaine non identifiées). Le manuel d'utilisation de décembre 1990 présente de façon détaillée les différentes phases de constitution du fichier : l'identification des nés en métropole au RNIPP ; le chiffrement et la saisie des dossiers complets de 1984 à 1988 ; l'appariement avec le fichier au quart du recensement de 1982 ; l'appariement avec le fichier du recensement de 1975 ; l'appariement avec la BDCOM (récupération de la géographie : agglomérations, tranches d'unités urbaines aux différents recensements) ; la consultation du RNIPP (récupération des décès) ; la création de variables de comparaison des recensements (repérage des changements de profession et des changements d'établissement).
[57] Oliver Sautory, « Plus de la moitié de la population a changé au moins une fois de commune en vingt ans », *Economie et statistiques*, n°209, Insee, avril 1988. Michel Isnard, communication sur la « mortalité différentielle » à la cinquième réunion du réseau CRICED, octobre 1989. Michel Isnard, communication sur « la mobilité en France de 1962 à 1982 » au congrès de l'ISI, septembre 1990. Guy Desplanques et Michel Isnard, communication sur la « mobilité géographique d'après l'EDP » au congrès européen de démographie, octobre 1991. Alain Chenu, « Itinéraires professionnels et mobilité sectorielle », *Cheminements professionnels et mobilités sociales*, La documentation française, octobre 1992.

électeurs. Pour ce qui concerne l'EDP, chaque dossier individuel est constitué d'une « chemise » et d'un « contenu ». Le « contenu » est constitué de l'ensemble des bulletins associés à la personne « EDP ». La « chemise », à l'image de la chemise carton utilisée auparavant, récapitule les informations utiles à l'identification des personnes et à la gestion des dossiers, en particulier, la date et le lieu de naissance, quelques variables biographiques issues des bulletins collectés jusqu'alors (dernier domicile connu, dernière catégorie socioprofessionnelle connue, date et lieu de naissance de la mère, date de naissance du dernier enfant…) et des variables de gestion. La chemise comporte également le numéro d'identification signifiant utilisé pour gérer l'EDP, constitué du premier NIR attribué à la personne et du siècle de naissance (variable NED/NEDTS)[58].

- *Création d'un poste de travail « gestionnaire EDP » pour gérer les dossiers et traiter l'identification des bulletins.* Ce poste de travail est relié à la fois aux dossiers individuels de l'EDP et à la base des répertoires des personnes physiques.

  Ce poste de travail permet d'abord de gérer les dossiers. En particulier, d'en créer de nouveaux pour insérer des nouveaux individus dans l'échantillon démographique permanent. Il permet aussi de supprimer un dossier d'une personne considérée intégrée à tort (suite à une erreur sur la date de naissance) ou de fusionner deux dossiers de personnes qu'on pensait distinctes et qui s'avèrent être identiques lors d'un examen ultérieur.

  Le poste de travail permet ensuite de traiter l'identification des bulletins d'état civil et des bulletins du recensement. En cas d'identification univoque, l'insertion du bulletin dans le dossier ou la création d'un nouveau dossier est automatique. C'est le cas dans environ neuf cas sur dix[59]. Dans le cas contraire, il y a litige. Les litiges sont repris par des gestionnaires. Pour les aider à retrouver la personne dans l'EDP ou au répertoire des personnes physiques, ils peuvent jouer sur les critères d'identification et consulter les informations biographiques présentes dans la « chemise »[60] .

  Pour l'état civil, les bulletins sont directement récupérés dans la BRPP. L'identité (nom, prénom) correspond à l'information brute issue de la saisie des bulletins statistiques ; les autres informations sont récupérées à l'issue des chaînes de traitement de l'état civil et bénéficient donc de ses contrôles de validité. L'identification des personnes est donc réalisée à partir d'une date et d'un lieu de naissance valides et des patronymes bruts.

  Pour le recensement, l'identification est réalisée à partir d'un fichier *ad hoc* qui comprend l'état civil complet de la personne, le numéro identifiant du bulletin individuel et des informations biographiques des autres personnes du ménage. L'ensemble des informations statistiques est récupéré ultérieurement en aval des chaînes de traitement du recensement, à l'aide du numéro identifiant du bulletin individuel.

  Enfin, le poste de travail permet de traiter les « incohérences internes » au dossier. Le gestionnaire peut ainsi modifier manuellement les informations de la chemise si celles qui y figurent, renseignées à partir des bulletins déjà intégrés, lui semblent erronées à l'occasion de l'intégration d'un nouveau bulletin[61]. Il peut aussi créer des bulletins « fantômes » dans le cas

---

[58] Le nom et le prénom ne figurent pas dans le chemise, sauf sous forme phonétique pour les personnes nées hors de France métropolitaine, parce que, pour ces personnes, l'EDP ne dispose que d'un NIR fictif.

[59] Dans le cas des bulletins statistiques de naissance de personnes nées un jour « EDP », la création d'un nouveau dossier est automatique s'il n'y a pas de dossier trop voisin déjà présent dans l'EDP. Dans le cas contraire, le bulletin de naissance passe en litige pour être traité par un gestionnaire. Pour les autres bulletins, il y a litige en cas d'absence de dossier suffisamment proche ou de l'existence de plusieurs voisins possibles.

[60] Pour traiter des litiges, le poste de travail fait apparaître la liste des dossiers voisins avec leurs principales caractéristiques, classés selon une « note de cohérence ». Il permet d'ajouter ou de réduire le nombre de caractéristiques utilisées pour définir le nombre de dossiers voisins. Il permet d'accéder au contenu du bulletin, notamment pour regarder les éventuels cas de noms ou prénom mal orthographiés. Il permet enfin une recherche « multi critère » sur la base des variables biographiques de la « chemise » en faisant abstraction des nom et prénoms.

[61] Seules les variables de la « chemise », utilisées pour l'identification, sont modifiables. Les variables associées aux bulletins ne sont pas modifiées mais un indicateur de gestion indique que le bulletin est erroné, sans préciser la nature de l'erreur. De ce fait, dans la base d'études, des variables analogues issues de sources différentes (un

où le parcours biographique décrit par les bulletins insérés du dossier suggèrent l'existence d'autres actes d'état civil non insérés dans le dossier (par exemple, un bulletin de naissance d'enfants suggérant que les parents sont déjà mariés). Le bulletin « fantôme » est alors rempli partiellement par le gestionnaire à partir des informations qu'il peut récupérer par enquête auprès de la mairie concernée.

- *Traitement des bulletins d'état civil, en aval des chaînes des statistiques d'état civil et du RNIPP.* L'échantillon démographique permanent bénéficie ainsi des traitements de codification et de vérification faits dans ces chaînes. Les redressements opérés par ces chaînes ne sont toutefois pas pris en compte. L'intégration des bulletins d'état civil dans l'EDP n'est effectuée qu'après la production des fichiers statistiques d'état civil, dans le cadre d'une campagne annuelle.

- *Arrêt de la saisie des bulletins du recensement pour les besoins propres à l'EDP*. Désormais, les informations du recensement sont récupérées en aval des chaînes de traitement statistique du recensement. Pour cela, les bulletins des ménages ayant au moins une personne née un jour « EDP » sont systématiquement inclus dans l'exploitation « lourde ». Les informations des bulletins individuels des personnes nées un jour « EDP » saisies spécifiquement pour les besoins de l'EDP sont réduites : on y trouve en particulier le nom et le prénom qui, avec la date et le lieu de naissance et quelques informations complémentaires permettent d'identifier les personnes pour retrouver celles qui sont déjà présentes dans l'EDP. Les informations du recensement aux différents niveaux (individu, famille, ménage, logement) sont récupérées dans un second temps, via les numéros d'identification des bulletins. A partir du recensement de 1990, les variables « brutes » de l'EDP ne correspondent donc plus à une saisie spécifique réalisée pour les besoins de l'EDP mais aux variables « brutes » des exploitations du recensement, avant traitement des anomalies, redressements et imputation de la non-réponse.

- *Arrêt des échanges de bulletins entre directions régionales* : la direction régionale collectant un bulletin « EDP » devient responsable de sa gestion, indépendamment du lieu de naissance de la personne.

En rythme de croisière, ce nouveau système doit permettre de livrer au moins une base d'études par an pour intégrer une année supplémentaire d'évènements d'état civil, au plus tard deux ans après l'année considérée.

Le projet d'informatisation, lancé en 1992, achève son déploiement en 1995. Pendant cette période de transition entre l'ancien mode de gestion et le nouveau, il faut chiffrer et saisir les 800 000 dossiers existants, intégrer les bulletins du recensement de 1990, intégrer le stock de bulletins d'état civil accumulés au cours de la période intercensitaire 1982-1989 et commencer à intégrer en continu les bulletins d'état civil des années plus récentes. Les moyens humains insuffisants obligent alors à des arbitrages qui conduisent à restreindre l'intégration de certains bulletins d'état civil à la moitié du champ de l'EDP :

- pour la récupération du stock de la période 1982-1989, les bulletins de mariage, de naissances d'enfants de personnes « EDP » et de reconnaissance ne sont intégrés par la plupart des directions régionales que pour les personnes nées les 1er et 4 octobre (les bulletins de naissance de personnes « EDP » sont toutefois intégrés pour les quatre jours de naissance) ;

- pour l'actualisation des informations d'état civil à compter de 1990, les bulletins d'état civil ne sont intégrés que pour les personnes nées les 1er et 4 octobre.

De son côté, l'introduction des bulletins du recensement de 1990 est effectuée pour l'essentiel avant la mise en œuvre des nouvelles applications. Une application *ad hoc* est donc utilisée selon des principes proches de la nouvelle application : quelques variables d'identification, notamment l'état civil complet et le numéro de bulletin individuel, sont saisies ; l'application réalise ensuite des contrôles de

---

bulletin de recensement et un bulletin d'état civil), peuvent proposer des informations divergentes pour le sexe, l'année de naissance ou le lieu de naissance.

cohérence et permet de vérifier l'existence du NIR dans la dernière version de l'EDP (base 1968-1982) ou du bulletin individuel dans le recensement. Une fois réalisée cette phase d'identification, les informations du recensement de 1990 sont intégrées.

Afin de « cibler sur sa vocation naturelle de fichier démographique », certaines variables des recensements de 1968 à 1982 présentes dans la base 1968-1982, jugées « surabondantes ou périphériques »[62] ne sont pas reprises dans la nouvelle base.

La première base d'études issue de la rénovation est livrée au deuxième trimestre 1995, avec les informations des recensements de 1968 à 1990 et les informations d'état civil de 1968 à 1989[63]. La base est ensuite actualisée régulièrement par intégration des nouvelles informations d'état civil. L'EDP commence alors à être davantage utilisé. Un dossier spécial d'*Économie et Statistiques* lui est consacré en octobre 1998 (n°316).

A la fin des années quatre-vingt-dix, l'Insee met en œuvre le projet d'alimentation informatisée du répertoire des personnes physiques (Airepp) qui permet aux mairies adhérentes de transmettre les données d'état civil sous forme télématique. Les traitements de la BRPP sont regroupés sur un nombre réduit de directions régionales sites. Le poste de gestion EDP n'est pas modifié ni les principes généraux de traitements.

- ***Le projet « EDP+ »***

Avec l'abandon du principe d'une collecte exhaustive espacée dans le temps au profit d'une collecte annuelle par échantillon, constitué de telle sorte que le recensement est basé sur le cumul de cinq années de collecte annuelle, la nouvelle méthode de recensement présentait deux difficultés pour l'EDP. D'abord, le nombre d'individus « EDP » recensés une année donnée diminuait fortement du fait de l'échantillonnage, diminuant la quantité d'information disponible à une date donnée (l'intersection d'une enquête annuelle de recensement avec l'EDP sur quatre jours de naissance est de l'ordre de 80 000 personnes). Ensuite, sur le plan informatique, le nombre de variables allait croitre très rapidement avec l'inclusion d'un nouveau lot de variables chaque année, avec des informations à blanc pour la majorité des individus. Une refonte de l'EDP s'avérait donc nécessaire pour restructurer la base et l'enrichir par des informations issues de nouvelles sources susceptibles d'apporter des informations analogues à celles issues du recensement, mais aussi de nouvelles informations.

Ce projet d'élargissement s'inscrit également dans le cadre des orientations proposées dans le rapport de Mylène Chaleix et Stefan Lollivier sur les « outils de suivi des trajectoires des personnes en matières sociales et d'emploi » qui préconise un développement des panels[64]. Le projet « EDP+ » en constitue la première phase ; le projet « EDP++ », la seconde.

Les premiers documents relatifs au projet « EDP+ » envisageaient différents aspects :
- La restructuration de la base d'études pour passer d'une table unique à une série articulée de tables par année de recensement et type de bulletins d'état civil ;
- L'adaptation des chaînes de production de la base d'études aux évolutions de la BRPP prévues dans le cadre de sa rénovation (projet BRPP2) - cette adaptation implique notamment le passage à un environnement Oracle et la réécriture de toutes les chaînes de traitement ;
- L'intégration des informations sur l'inscription électorale ;
- La multiplication de la taille d'échantillon par quatre pour augmenter le nombre d'individus « EDP » présents dans une enquête annuelle de recensement donnée (notamment dans la

[62] Les termes sont de Dominique Rouault, dans son article du *Courrier des statistiques.*
[63] La base est une table SAS de 1304 variables et 779 043 observations, correspondant à 664 591 personnes nées en France métropolitaine et 114 452 personnes nées hors de métropole. Les informations géographiques des recensements sont fournies selon le référentiel de la date de collecte et dans le dernier référentiel disponible. La catégorie socioprofessionnelle est proposée selon la nomenclature de 1982 pour les quatre recensements.
[64] Rapport n°98/B010 de juin 2004. Parmi ses recommandations : *« Recommandation 1a : créer un tel panel à partir de l'EDP élargi (EDP++) à la fois en terme de taille, mais surtout en l'enrichissant à partir de sources administratives (fiscales et sociales) tout en respectant l'anonymat des individus et les règles de confidentialité. ».*

perspective de la réalisation de travaux qualité sur les enquêtes annuelles de recensement et de l'utilisation de l'EDP comme base de sondage) ;

- L'élargissement du champ géographique aux départements d'outre-mer à partir des évènements relatifs à l'année 2004.

En 2011, une base d'études d'attente est livrée, limitée aux personnes nées les quatre premiers jours d'octobre. Conforme au schéma habituel pour les données d'état civil et les recensements généraux de population, elle propose également une série de tables distinctes pour les enquêtes annuelles de recensement 2004 à 2009.

En 2012, une première base d'études « EDP+ » produite à partir des nouvelles chaînes est livrée, portant sur les seize jours de naissance et l'ensemble de la France et incluant les informations du fichier électoral.

- ***La « migration » vers la BRPP rénovée***

La mise en œuvre de la BRPP rénovée a contraint à modifier l'architecture informatique de production de la base d'études de l'EDP. Elle a aussi nécessité une phase d'initialisation ou « migration » des individus, au cours de laquelle les individus EDP devaient être ré-identifiés. Cette migration, effectuée en 2010, n'a toutefois été réalisée que sous certaines conditions :

- *Pour les personnes nées en métropole selon l'EDP* : elles sont censées être inscrites au répertoire des personnes physiques ; la migration n'a donc été réalisée que lorsqu'elles ont été effectivement retrouvées au répertoire. En pratique, une extraction des personnes nées un jour « EDP » présentes dans la BRPP rénovée a été réalisée ; leur identifiant théorique dans l'EDP a été reconstruit à partir de leur NIR et de leur siècle de naissance (sur le modèle de la variable Ned) ; seules les personnes effectivement retrouvées dans l'EDP ont ensuite été migrées. De ce fait, les personnes nées avant 1891 n'ont pas été migrées (elles sont absentes de la BRPP rénovée). Les personnes dont le NIR a changé au cours du temps ne l'ont pas été non plus (la BRPP rénovée s'appuie sur le NIR actif tandis que la variable Ned de l'EDP est construite à partir du premier NIR attribué)[65]. Dans le cadre de l'initialisation réalisée par la BRPP, aucune recherche particulière n'a été réalisée pour voir si des personnes de l'EDP non migrées pouvaient correspondre à des personnes nées un jour « EDP » dans la BRPP2 qui n'avaient pas été retrouvées dans l'EDP lors du rapprochement automatique[66].

- *Pour les personnes nées hors métropole* : l'EDP leur a attribué un pseudo NIR, indépendamment du vrai NIR qu'une partie d'entre elles pouvaient avoir à la « section hors métropole ». La migration n'a donc été réalisée que si les informations disponibles sur le sexe, la date et le lieu de naissance permettait de retrouver la personne de façon univoque au répertoire. Il n'y a pas eu de recherche complémentaire. En particulier, la variable Soundex déduite du nom et du prénom conservés dans l'EDP pour les personnes nées hors métropole n'a pas été utilisée.

Du fait de l'architecture informatique retenue, les informations statistiques accumulées sur les personnes « non migrées » ne peuvent pas être intégrées avec les informations migrées dans la nouvelle base de gestion. Ces trajectoires « non migrées » ont toutefois été réintégrés par la mise en place d'une chaîne de traitements *ad hoc* en marge du cadre normal, grâce à des développements informatiques spécifiques. Cependant, la prolongation des trajectoires « non migrées » ne sera plus possible. Si de nouvelles informations arrivent sur des individus « non migrés », un nouvel individu sera créé amorçant une nouvelle trajectoire.

[65] Par ailleurs, les personnes nées en Corse quand leur identifiant dans l'EDP (variable Ned) utilisait les codes département « 2A » ou « 2B » n'avaient pas été migrés initialement mais ce cas a pu être traité rapidement par un correctif apporté dans les chaînes de production de l'EDP.

[66] Des travaux exploratoires réalisés ensuite par la division des enquêtes et études démographiques suggèrent que quelques centaines d'individus pourraient être concernées.

Selon des travaux réalisés par la division des enquêtes et études démographiques, 43 % des personnes nées à l'étranger n'ont pas été migrées, soit 111 780 personnes. La plupart a intégré l'EDP à l'occasion des recensements de 1968 ou 1975 mais 23 % des personnes concernées ont été recensées en 1999 et 13 % dans une enquête annuelle de recensement. Seulement 9 % disposent d'un bulletin de décès[67]. Par ailleurs, 30 000 personnes nées en métropole n'auraient pas été migrées, dont 40 % nées avant 1891.

Le tableau de la page suivante (figure 5) propose des comparaisons d'effectifs entre une base d'études antérieure à la rénovation de la BRPP[68], une base de travail produite après la « migration » (dite la base « de recette ») et la base d'études millésimée « 2012 ». Ces deux dernières ont donc été produites après la « migration » vers la BRPP rénovée. La base « de recette » n'intègre pas les personnes « non migrées ». La base d'études millésimée « 2012 » les réintègre. Les comptages sont effectués sur les personnes EDP présentes aux différents recensements généraux de population. Les comparaisons avec la base « de recette de l'EDP+ » permettent de confirmer que la question des « non migrées » se concentre sur les personnes nées hors métropole. Les comparaisons avec la base d'études millésimée « 2012 » montrent que la chaîne de traitements *ad hoc* a bien permis de réintégrer les trajectoires passés des « non migrées »[69].

La « migration » dans le nouvel environnement de la BRPP n'a pas concerné que les personnes. Elle a aussi concerné les évènements. En particulier, les évènements d'état civil qui alimentent la base d'études de l'EDP sont désormais conservés de façon définitive dans la sphère de la BRPP. Les évènements d'état civil relatifs aux personnes « migrées » ont donc été intégrés dans la BRPP. Pour cela, il fallait que leur structure ou que leur existence soit compatible avec les bulletins d'état civil gérés par la BRPP rénovée. De ce fait, les bulletins de reconnaissance et les bulletins de mariage légitimant une personne « EDP » n'ont pas été migrés et ne figurent donc pas dans la nouvelle base d'études. Certaines informations des bulletins de mariages n'ont pas non plus été récupérées (elles devraient l'être par la suite).

---

[67] Pour les personnes décédées, il serait possible de réintégrer l'ensemble de leur trajectoire dans l'EDP sans risque que la personne réapparaisse à l'avenir avec un autre identifiant.

[68] Pour des raisons pratiques, le millésime 2003 a été utilisée. Il ne s'agit donc pas de la dernière base d'études produite avec les anciennes chaînes de la BRPP, mais cela n'est pas gênant dans la mesure où les comptages présentés sont effectués pour les recensements généraux, dont le dernier remonte à 1999.

[69] Les effectifs ne sont toutefois pas strictement identiques. D'une part, parce qu'à la constitution d'une nouvelle base, une personne considérée comme EDP peut ne plus l'être une autre fois en raison de la prise en compte de nouvelles informations.

Figure 5 - Comparaison des effectifs contenus dans la base d'études de 2012, des effectifs dans la base EDP+ en recette et des effectifs dans la base d'études millésimée 2003

Sexe	Lieu de naissance	Effectifs dans la base études 2003	Effectifs dans la base de recette de l'EDP+	Effectifs dans la base d'études 2012	Ecart sur les effectifs avant réintégration des non migrés	Taux de récupération estimé avant réintégration des non migrés	Ecart sur les effectifs après réintégration des non migrés	**Taux de récupération estimé**
Recensement de 1968								
Hommes	Métropole	206 991	198 332	206 981	-8 659	95,8%	-10	**100,0%**
Femmes	Métropole	223 507	211 731	223 482	-11 776	94,7%	-25	**100,0%**
Hommes	Hors métropole	28 317	10 926	28 265	-17 391	38,6%	-52	**99,8%**
Femmes	Hors métropole	24 704	9 608	24 671	-15 096	38,9%	-33	**99,9%**
		483 519	430 597	483 399	-52 922	89,1%	-120	**100,0%**
Recensement de 1975								
Hommes	Métropole	224 377	222 935	224 368	-1 442	99,4%	-9	**100,0%**
Femmes	Métropole	241 585	237 814	241 568	-3 771	98,4%	-17	**100,0%**
Hommes	Hors métropole	31 264	13 552	31 207	-17 712	43,3%	-57	**99,8%**
Femmes	Hors métropole	28 067	11 783	28 024	-16 284	42,0%	-43	**99,8%**
		525 293	486 084	525 167	-39 209	92,5%	-126	**100,0%**
Recensement de 1982								
	Métropole	489 830	488 228	489 801	-33 274	99,7%	-29	**100,0%**
	Hors métropole	62 770	29 496	62 647	-33 274	47,0%	-123	**99,8%**
		552 600	517 724	552 448	-105 757	93,7%	-152	**100,0%**
Recensement de 1990								
Hommes	Métropole	251 602	250 361	250 951	-1 241	99,5%	-651	**99,7%**
Femmes	Métropole	270 353	268 966	269 693	-1 387	99,5%	-660	**99,8%**
Hommes	Hors métropole	31 929	17 564	31 797	-14 365	55,0%	-132	**99,6%**
Femmes	Hors métropole	32 007	16 148	31 889	-15 859	50,5%	-118	**99,6%**
		585 891	553 039	584 330	-32 852	94,4%	-1 561	**99,7%**
Recensement de 1999								
Hommes	Métropole, DOM, TOM	266 109	265 318	266 115	-791	99,7%	6	**100,0%**
Femmes	Métropole, DOM, TOM	285 813	284 786	285 814	-1 027	99,6%	1	**100,0%**
Hommes	Etranger	30 041	21 614	29 991	-8 427	71,9%	-50	**99,8%**
Femmes	Etranger	31 334	19 832	31 284	-11 502	63,3%	-50	**99,8%**
		613 297	591 550	613 204	-21 747	96,5%	-93	**100,0%**
Variables utilisées pour les comptages :								
Les variables "sexe" et "indicateur de lieu de naissance" associées aux différents recensements								
S68, ILN68, S75, ILN75, ILN82, S90, ILN90, S99, ILN99.								

- ***Le projet « EDP++ »***

Le projet « EDP++ » constitue une nouvelle étape dans l'enrichissement de l'échantillon démographique permanent. Il consiste à intégrer des informations issues de nouvelles sources administratives externes à l'échantillon démographique permanent, comme les DADS ou les déclarations fiscales. Ces sources, qui comportent certaines informations sociodémographiques de même nature ou sur des thèmes analogues à ceux couverts par le recensement, présentent l'avantage de l'exhaustivité sur leur champ. Elles pourront donc permettre de réaliser des études en profitant pleinement de la taille de l'échantillon démographique permanent, là où le recensement, qui s'effectue désormais par échantillon, ne le permet plus, par exemple sur la mobilité géographique ou les trajectoires socioprofessionnelles.

Dans ce nouveau cadre, les informations issues du recensement gardent cependant leur intérêt, par exemple pour définir des cohortes à partir d'informations absentes des sources administratives, comme le diplôme, cohortes dont on pourra suivre ensuite l'évolution à partir des informations issues des sources administratives. Elles permettront aussi d'étudier les éventuels écarts entre les informations déclarées au recensement et les informations statistiques issues de données administratives.

Le cadre juridique

L'échantillon démographique permanent a été créé avant le vote de la loi Informatique et Libertés. Après le vote de cette loi, la démarche de régularisation a été entamée et a abouti à deux textes réglementaires qui autorisent explicitement et encadrent l'échantillon démographique permanent :

- Le **décret n° 84-393 du 23 mai 1984** autorise l'utilisation du répertoire d'identification des personnes physiques pour le traitement automatisé de l'échantillon démographique permanent ;

- L'**arrêté du 23 mai 1984** modifié « définit » succinctement l'échantillon démographique permanent. L'article 1 précise que la « seule finalité » est « l' élaboration de statistiques démographiques et sociales ». L'article 2 précise les personnes concernées (celles « nées du 2 au 5 janvier, du 1er au 4 avril, du 1er au 4 juillet et du 1er au 4 octobre de chaque année »). Il indique également que les informations prises en compte sont « issues des recensements de la population successifs, des bulletins statistiques de l'état civil et du fichier électoral », sans entrer dans le détail. L'article 4 initial, abrogé en 2010, interdisait de « communiquer à quiconque des informations nominatives issues de l'échantillon démographique permanent ».

L'arrêté du 23 mai 1984 a été modifié à deux reprises :

- l'**arrêté modificatif du 21 décembre 2006**, ajoute douze jours de naissance aux quatre jours de naissance initiaux, permettant le quadruplement de la taille de l'échantillon[70].

- l'**arrêté modificatif du 20 janvier 2010[71]** modifie les règles d'accès à l'échantillon démographique permanent. Il modifie l'article 3 et abroge l'article 4 initial, qui interdisait de donner accès aux informations nominatives de l'EDP à des tiers externes à l'INSEE.

  Jusqu'alors, le caractère indirectement nominatif des informations de l'EDP du fait de leur richesse et du degré de détail (par exemple par croisement du sexe, de la date et de la commune de naissance) obligeait l'INSEE à signer avec les organismes de recherche des conventions de mise à disposition des chercheurs qui souhaitaient travailler sur l'EDP. Désormais, l'accès est possible selon le régime commun mis en place par le décret n°2009-318 du 20 mars 2009 relatif au Conseil national de l'information statistique et au comité du secret statistique : les renseignements individuels permettant l'identification des personnes peuvent être communiqués « à des fins de statistique publique ou de recherche scientifique ou historique, sur décision de l'administration des archives prise après avis du comité du secret statistique et accord de l'Institut national de la statistique et des études économiques ». Dans ce cas, les chercheurs accèdent aux données par l'intermédiaire du centre d'accès sécurisé distant (CASD) mis en place par le Groupe des écoles nationales d'économie et statistique (GENES). Sa technologie permet à des utilisateurs clairement identifiés d'effectuer des traitements sur des données individuelles tout en empêchant la sortie de données ne garantissant pas le secret statistique.

Les textes réglementaires de 1984 ont été précédés d'un avis favorable motivé de la CNIL[72] :

> *Le Ministre de l'Économie et des Finances a adressé, le 6 octobre 1980, une déclaration du traitement d'informations nominatives dénommé "Échantillon démographique permanent" mis en œuvre par l'Institut National de la Statistique et des Études Économiques, accompagnée*

[70] L'avis favorable de la CNIL en date du 14 novembre 2006 rappelle explicitement les dispositions de la loi Informatique et Liberté stipulant que les personnes auprès desquelles sont recueillies les données à caractère personnel doivent être informées de l'identité du responsable du traitement, de ses finalités et des destinataires.
[71] Saisine de la CNIL en date du 5 novembre 2009.
[72] Délibération n°83-26 du 12 avril 1983 portant avis sur les traitements automatisés d'informations nominatives créés à l'INSEE à des fins d'élaboration statistique à partir d'un échantillon permanent de population dit "Échantillon démographique permanent".

d'une demande d'avis relative au projet de décret l'autorisant à utiliser le répertoire national d'identification des personnes physiques, en vue d'effectuer ce traitement.

[…]

La Commission Nationale de l'Informatique et des Libertés émet l'avis suivant :

Sur la Procédure :

Considérant qu'au vu de la déclaration du traitement de l'échantillon démographique permanent il apparaît opportun qu'il soit procédé à un examen particulier de ce traitement et qu'intervienne un arrêté ministériel le réglementant, la Commission décide, outre son avis sur le projet de décret autorisant l'utilisation du répertoire national d'identification des personnes physiques pour effectuer ce traitement, d'émettre un avis motivé sur le projet d'arrêté le réglementant complémentairement.

Sur le fond :

La Commission prend acte de ce que :

- le traitement a pour seule finalité l'élaboration de statistiques démographiques et sociales ;
- les personnes concernées sont celles nées entre le 1er et le 4 octobre de chaque année ;
- les informations traitées peuvent provenir des recensements généraux successifs de la population, des bulletins d'état civil, du fichier électoral et du répertoire national d'identification des personnes physiques, ces deux derniers étant réglementés par ailleurs ;
- en ce qui concerne le rapprochement avec le répertoire national d'identification des personnes physiques, il ne s'agit, à partir de l'identité des personnes ainsi que de leurs date et lieu de naissance figurant dans les diverses sources, que d'obtenir le numéro d'inscription au répertoire afin de l'utiliser comme identifiant permanent dans le fichier informatisé constitué ;
- une fois cette opération réalisée, les nom et prénoms des personnes concernées ne sont pas conservés dans le fichier informatisé ;
- l'Institut National de la Statistique et des Etudes Economiques (INSEE) est seul destinataire des informations nominatives susvisées ;
- par suite, l'INSEE ne peut procéder qu'à des cessions d'informations ne permettant pas l'identification directe ou indirecte des personnes concernées ;
- sauf dispositions législatives contraires, la conservation des informations nominatives est soumise aux dispositions de la loi n° 79-18 du 3 janvier 1979 sur les archives ;
- le droit d'accès prévu à l'article 34 s'exerce auprès des directions et services régionaux de l'INSEE ;
- en cas d'erreur, l'INSEE procède aux rectifications demandées en application de l'article 36 sur le fichier du recensement et, éventuellement, sur le répertoire national d'identification des personnes physiques ainsi que le fichier électoral ;
- outre les dispositions usuelles prises par l'INSEE pour assurer la sécurité des traitements et des informations, des dispositions complémentaires sont prises dans le cas présent visant, dans chaque direction régionale, à placer sous la responsabilité d'un agent nommément désigné, les dossiers concernés, placés sous clés dans un local particulier ;

EMET, dans ces conditions, un AVIS FAVORABLE aux projets de décret et d'arrêté réglementant le traitement.

Enfin, la récupération des informations nominatives du recensement pour les besoins de l'EDP et leur traitement sont également abordés dans l'**arrêté du 12 janvier 2004**[73] (cet arrêté autorise la mise en œuvre des phases "saisie et exploitation des données collectées" et "contrôle de la cohérence des réponses aux enquêtes" du traitement "Recensement de la population"). Son article 5 stipule ainsi que :

[73] Modifié par l'arrêté 2007-04-30 art. 1 JORF 16 mai 2007.

« I. - Pour chaque personne faisant partie de l'échantillon démographique permanent (EDP), les données suivantes, à savoir ses nom, prénoms, sexe, date et lieu de naissance ainsi que le code à barres mentionné à l'article 2 du présent arrêté, font l'objet des opérations suivantes :
a) La création d'un fichier de saisie reprenant les données concernées à partir des images issues de la lecture automatisée des questionnaires ;
b) Le contrôle sur échantillon de la saisie des données donnant lieu à la création d'une base d'images de l'échantillon et du fichier-échantillon EDP qui lui correspond.

II. - Ce fichier de saisie est utilisé par l'INSEE pour mettre à jour l'échantillon démographique permanent par les données du fichier de saisie défini à l'article 2[74]*. Il est détruit au plus tard à la fin de l'année suivant celle de sa réception définitive par l'INSEE.*

III. - La base d'images de l'échantillon et le fichier-échantillon EDP qui lui correspond sont détruits par l'INSEE dans le délai mentionné à l'article 7[75]*, sous réserve des dispositions résultant de l'application de l'article 8 du présent arrêté.»*

---

[74] Les données de saisie listées dans l'article 2 correspondent à l'ensemble des réponses aux questionnaires du recensement, à l'exclusion du nom, du prénom. L'adresse et le numéro Code à barres sont inclus, ainsi que les variables externes ajoutées lors de l'exploitation complémentaire : l'activité économique, la catégorie juridique de l'employeur, la tranche d'effectifs salariés et la localisation de l'établissement employeur ou de l'établissement dirigé.

[75] Le prestataire de saisie détruit ces fichiers après réception à l'Insee des fichiers de saisie définitifs.

La documentation

🕮 Les manuels de l'utilisateur présentent de façon détaillée le contenu de la base d'études. Plusieurs versions successives ont été rédigées jusqu'en juillet 2003 sur les bases d'études antérieures au projet « EDP+ ». Depuis la mise en œuvre de « EDP+ », des nouveaux manuels sont également disponibles.
🕮 « Les sources des statistiques de l'état civil », 2011, dans la rubrique « Sources et Méthodes » sur www.insee.fr
🕮 Christine Couet, « L'échantillon démographique permanent de l'Insee », *Courrier des statistiques*, n°117-119, 2006.
🕮 Dominique Rouault, « L'échantillon démographique permanent a pris un coup de jeune : gestion informatisée et nouvelle base d'études », *Courrier des statistiques*, n° 73, mars 1995
🕮 Olivier Sautory, « L'échantillon démographique permanent de l'Insee », *Courrier des statistiques*, n° 41, janvier 1987.

Quelques exemples d'utilisation

Fécondité

🕮 Olivia Ekert-Jaffe., Heather Joshi et ali, « Fécondité, calendrier des naissances et milieu social en France et en Grande-Bretagne : politiques sociales et polarisation socioprofessionnelle », *Population* n°3, Ined, mai - juin 2002.
🕮 Isabelle Robert-Bobée, Michael Rendall et ali, « Âge au premier enfant et niveau d'études : une analyse comparée entre la France, la Grande-Bretagne et la Norvège », *Données Sociales*, Insee, 2006.

Mortalité différentielle

🕮 Michel Isnard, « Mortalité différentielle », communication à la 5e réunion du réseau CICRED, octobre 1989 (Paris).
🕮 Annie Mesrine, « La surmortalité des chômeurs : un effet catalyseur du chômage ? », *Économie et Statistiques,* n°334, Insee, 2000.
🕮 Christian Monteil et Isabelle Robert-Bobée, « Quelles évolutions des différentiels sociaux de mortalité pour les hommes et les femmes ? », *Document de travail*, n°F0506, Insee, 2006.
🕮 Emmanuelle Cambois, « Carees and mortality in France : Evidence on how far occupational mobility predicts differential risks », *Social Science and Medecine,* n°58, 2006.
🕮 Nathalie Blanpain, « L'espérance de vie s'accroît, les inégalités sociales face à la mort demeurent », *Insee Première,* n°1372, Insee, octobre 2011.
🕮 Emmanuelle Cambois et Caroline Laborde, « Mobilité socioprofessionnelle et mortalité en France. Des liens qui se confirment pour les hommes et qui s'affirment pour les femmes », *Population*, 66(2), 2011.

Autres travaux sur la mortalité

🕮 Xavier Niel, « Les facteurs explicatifs de la mortalité infantile en France et leur évolution récente. L'apport de l'échantillon démographique permanent. », *Document de travail*, n°F1106, Insee, juin 2011.

Mobilité résidentielle

🕮 Olivier Sautory, « Près de la moitié de la population a changé au moins une fois de commune en 20 ans », *Économie et statistique*, n°209, Insee, avril 1988.
🕮 Guy Desplanques et Michel Isnard, « Mobilité géographique d'après l'EDP », communication au congrès européen de démographie, octobre 1991 (Paris).
🕮 Ray Hall , Philip E. Ogden, « La mobilité des personnes seules en France et en Grande-Bretagne », *Economie et statistique*, n°316, Insee, 1998.

🕮 Christine Couet, « La mobilité résidentielle des adultes : existe-t-il des parcours-type? », *France Portrait Social*, Insee, 2006.

🕮 Matthieu Solignac, *La mobilité en mouvement, essais sur le transport, la mobilité et les disparités spatiales*, Thèse pour le doctorat de sciences économiques soutenue le 10 décembre 2013 à l'Université de Paris I Panthéon Sorbonne.

Mobilité socioprofessionnelle

🕮 Alain Chenu, « Itinéraires professionnels et mobilité sectorielle » in *Cheminements professionnels et mobilités sociales,* La Documentation française, octobre 1992.

🕮 Anne-Françoise Molinié, « Des secteurs et des âges », *Population*, 48e année, n°6, Ined, 1993.

🕮 Alain Chenu, « De recensement en recensement, le devenir professionnel des ouvriers et employés », *Economie et Statistique*, n°316, Insee, 1998.

🕮 Dominique Rouault et Olivier Galland, « Devenir cadre dès trente ans : une approche longitudinale de la mobilité sociale », *Economie et Statistique*, n°316, Insee, 1998.

🕮 Chantal Brutel, Maryse Jegou et Carole Rieu, « La mobilité géographique et la promotion professionnelle des salariés : une analyse par aire urbaine », *Économie et statistique,* n°336, Insee, juin 2000.

🕮 Cécile Détang-Dessendre, Virginie Piguet et Bertrand Schmitt, « les déterminants micro-économiques des migrations urbain-rural : leur variabilité en fonction de la position dans le cycle de vie », *Population*, 57, Ined, 2002.

🕮 Virginie Christel, « Trajectoires résidentielles des personnes âgées », *Données sociales édition 2006*, Insee, 2006.

🕮 Marie-Paule Couto, « L'intégration socio-économique des pieds-noirs en France métropolitaine : le lien de citoyenneté à l'épreuve », Revue européenne des migrations internationales, n°29(3), septembre 2013.

Emploi, lien formation-emploi

🕮 Olivier Galland et Dominique Rouault, « Des études supérieures inégalement rentables selon les milieux sociaux », *Insee première,* n°469, Insee, juillet 1996

🕮 Anne-Françoise Molinié, « Déclin et renouvellement de la main d'œuvre industriel », *Economie et Statistique*, n°316, Insee, 1998.

Insertion des immigrés

🕮 Jean-Luc Richard, « Unemployment among young people of foreign origin in France : Ways of measuring discrimination», p. 1001-135, *the Way to a Multicultural Society, Sw*iss Federal Statistical office (ed.), Bern, 1997.

🕮 Jean-Luc Richard, « Rester en France, devenir français, voter : trois étapes de l'intégration des enfants d'immigrés », *Économie et Statistique,* n°316-317, Insee, juin-juillet 1998

🕮 Jean-Luc Richard, « Une approche de la discrimination sur le marché du travail, les jeunes adultes issus de l'immigration étrangère en France *», Revue européenne des migrations internationales*, n°16-3, 2000.

🕮 Jean-Luc Richard, *Partir ou rester ? Destinée des jeunes issus de l'immigration*, PUF, Paris, 2004.

🕮 Fanny Mikol et Chloé Tavan, « La mobilité professionnelle des employés et ouvriers immigrés », *Données Sociales*, Insee, 2006.

🕮 Denis Fougère et Mirna Safi, « L'acquisition de la nationalité française : quels effets sur l'accès à l'emploi des immigrés? », *France Portrait Social 2005 - 2006*, Insee, 2006.

🕮 Mirna Safi, « Inter-mariage et intégration : les disparités des taux d'exogamie des immigrés en France », *Population*, n°63*,* Ined, 2008.

🕮 Rahsaan Maxwell, *Pour en finir avec un faux débat : les statistiques ethniques*, En Temps réel, 2009.

Utilisations régionales

🕮 Nadine Laroche, « Femme échange moins volontiers travail contre bébé », *Regards sur l'Ile- de-France*, n° 32, Insee Île-de-France, Juin 1996.
🕮 Dominique Roussel, « Partir, revenir », *Économie lorraine,* n°156, Insee Lorraine, septembre 1996.
🕮 C. Devos, « Des parcours professionnels mouvementés de 1968 à 1980 », *Profils Nord - Pas-de-Calais,* n° 2, Insee Nord-Pas-de-Calais, février 1997.

Participation électorale[76]

🕮 Jean Morin, « Un français sur dix ne s'inscrit pas sur les listes électorales », *Économie et statistique*, n°152, Insee, février 1983.
🕮 Jean Morin, « La participation électorale dans les grandes villes aux élections municipales de 1983 », *Économie et statistique*, n°165, Insee, avril 1984.
🕮 François Héran, « Les intermittences du vote : un bilan de la participation de 1995 à 1997 », *Insee Première,* n° 546, Insee, septembre 1997.
🕮 François Clanché, « La participation électorale au printemps 2002 : de plus en plus de votants intermittants », *Insee Première* n° 877, Insee, janvier 2003.
🕮 Stéphane Jugnot, « La participation électorale en 2007 : la mémoire de 2002 », *Insee Première,* n° 1169, Insee, décembre 2007.
🕮 Stéphane Jugnot, Nicolas Frémeaux « Les enfants des baby-boomers votent par intermittence, surtout quand ils sont peu diplômés », *France, portrait social,* Edition 2010, Insee.
🕮 Xavier Niel et Liliane Lincot, « L'inscription et la participation électorale en 2012 : qui est inscrit et qui vote », *Insee Première* n° 1411, Insee, septembre 2012.

Causes de mortalité[77]

🕮 Béatrice Geoffroy-Perez, « Analyse de la mortalité et des causes de décès par secteur d'activité de 1968 à 1999 à partir de l'échantillon démographique permanent », rapport « Cosmop », InVS, septembre 2006.
🕮 Gwen Menvielle, Annette Leclerc, Jean-François Chastang et Danièle Luce (groupe EDISC), « Social inequalities in breast cancer mortality among French women: disappearing educational disparities from 1968 to 1996 », *British Journal of Cancer*, 94, 2006.
🕮 Gwen Menvielle, Jean-François Chastang, Danièle Luce et Annette Leclerc (groupe EDISC), « Évolution temporelle des inégalités sociales de mortalité en France entre 1968 et 1996. Étude en fonction du niveau d'études par cause de décès », *Revue d'Épidémiologie et de santé publique*, 2007, volume 55, p97-105.
🕮 Marie-Josèphe Saurel-Cubizolles, Jean-François Chastang, Gwenn Menvielle, Annette Leclerc et Danièle Luce (groupe EDISC), « Social inequalities in mortality by cause of death in women and men in France », *Journal of Epidemiology and community health*, n° 63, 3, 2009.
🕮 Christine Cohidon, Gaëlle Santin, Béatrice Geoffroy-Perez et Ellen Imbernon, « Suicide et activité professionnelle en France », *Revue d'Épidémiologie et de Santé Publique*, Vol. 58/2, 2010.

---

[76] Travaux utilisant l'EDP enrichi par l'inscription électorale comme base de sondage pour des enquêtes sur la participation électorale.
[77] Travaux réalisés à partir de l'EDP enrichi par les causes de mortalité dans le cadre du projet « EDISC » (Évolution des inégalités sociales par causes médicales de décès), piloté par l'Inserm, et du projet « COSMOP » (Cohorte pour la surveillance de la mortalité par profession), piloté par l'Institut national de veille sanitaire.

ANNEXE 1 - Description de la nouvelle base d'études EDP+

La description de la nouvelle base d'études « EDP+ » présentée ici est synthétique. Une version actualisée ainsi que la liste détaillée des variables et de leurs modalités est disponible dans le manuel de l'utilisateur de la base d'études.

Table : INDIVIDU	**Unité observée :** L'individu « EDP »
Source : variables de gestion (variables calculées)	
Cette table comprend quelques caractéristiques principales de la personne dès lors que les bulletins intégrés permettent d'en disposer : sexe, âge, date et lieu de naissance, date et lieu de décès, lieu et date de naissance des parents, ainsi que des indicateurs d'évènements le concernant qui sont présents dans la base d'étude : acte de naissance, acte de décès, nombre d'enfants repérés dans les actes d'état civil, nombre de mariages repérés dans les actes d'état civil, nombre de bulletins à chaque recensement et enquête annuelle de recensement. **Cette table peut faciliter la sélection des personnes que l'on souhaite étudier (personnes nées telles années ; repérées dans telle source...).**	

Table : NAISSANCE	**Unité observée :** L'individu « EDP »
Source : ***État civil*** (bulletin de naissance, transcription d'un jugement déclaratif de naissance, transcription d'un jugement d'adoption plénière).	
Informations relatives à la naissance de l'individu « EDP », notamment : sexe, date et lieu de naissance, indicateur de naissance multiple ; date de naissance, lieu de naissance, nationalité, catégorie socioprofessionnelle de chacun des parents. ***Remarques :*** 1) Les mort-nés nés un jour EDP ne sont pas intégrés comme individu « EDP ». 2) En cas d'adoption plénière de l'enfant « EDP » par un tiers, les informations sur ses parents initiaux sont supprimées et remplacées par celles relatives aux parents adoptants. 3) Les informations sur la légitimation de l'enfant « EDP » par mariage de ses parents ne sont pas intégrées dans la nouvelle base d'études (les évènements enregistrés dans le passé, auparavant présents dans la base d'études, n'ont pas été migrés dans BRPP2). 4) Les informations sur la reconnaissance de l'enfant « EDP » par l'un de ses parents ne sont pas intégrées dans la nouvelle base d'études les évènements enregistrés dans le passé, auparavant présents dans la base d'études, n'ont pas été migrés dans BRPP2).	

Table : MARIAGE	**Unité observée :** L'individu « EDP »
Source : ***État civil*** (bulletin de mariage).	
Informations relatives au mariage de l'individu « EDP » : date et lieu de mariage ; date et lieu de naissance, situation matrimoniale antérieur et nationalité de chaque époux, lieu de résidence des époux. ***Remarques :*** 1) Seule la date de naissance des époux et la date de mariage ont été récupérées lors de la migration. Pour les évènements antérieurs à 2010, plusieurs informations disponibles auparavant dans l'EDP n'ont pas été récupérées immédiatement suite à un défaut de migration : le lieu de mariage, le lieu de naissance, la nationalité et la situation matrimoniale antérieure, le lieu de résidence des époux, leur catégorie socioprofessionnelle (pour les évènements antérieurs à 1998) ou le nombre d'enfants légitimés par le mariage. Les informations seront récupérées dans la base « 2013 ». 2) Le nombre de mariages intégrables n'est plus borné. 3) Une petite partie des bulletins de mariages célébrés en France ne sont pas transmis à l'Insee (l'Institut procède de ce fait à un redressement systématique du nombre total de mariages dans ses statistiques d'état-civil ; il n'est évidemment pas possible d'effectuer ce redressement au niveau individuel).	

Table : DESCENDANCE	**Unité observée :** Les enfants d'individus « EDP »
Source : ***État civil*** (bulletin de naissance, transcription d'un jugement déclaratif de naissance, transcription d'un jugement d'adoption plénière, enfant sans vie).	
Informations relatives à la naissance des enfants « EDP », notamment : sexe, date et lieu de naissance, indicateur de naissance multiple ; date de naissance, lieu de naissance, nationalité, catégorie socioprofessionnelle de chacun des parents (dont l'individu « EDP »). ***Remarques :*** 1) Les enfants nés sous X de mères « EDP » ne sont pas pris en compte. 2) Les enfants morts-nés déclarés à l'état-civil sont pris en compte. 3) Le nombre d'enfants associés à l'individu EDP n'est plus borné à 12. 4) Les reconnaissances par des individus « EDP » ne sont pas intégrées dans la nouvelle base d'études (ces événement ont cependant été migrés dans BRPP2).	

Table : DECES	**Unité observée :** L'individu « EDP »
Source : ***État civil*** (bulletin de décès, transcription d'un jugement déclaratif de décès, transcription d'un jugement déclaratif d'absence), ***RNIPP***	
Informations relatives au décès des individus « EDP », notamment : date et lieu de décès (commune et nature précise du lieu : domicile, hôpital, etc.) ; sexe, date et lieu de naissance, situation matrimoniale, catégorie socioprofessionnelle, nationalité et lieu de résidence du défunt.	

Table : FE	**Unité observée :** L'inscription d'un individu « EDP » dans une commune donnée à une date d'inscription donnée.
Source : ***Fichier général des électeurs***	
Historique des inscriptions des individus « EDP » dans le fichier général des électeurs : commune d'inscription, date d'inscription, état actif ou non de l'inscription et type de liste (« principal » ou « complémentaire »). ***<u>Remarques</u> :*** 1) Seules les inscriptions sur les listes électorales en France sont prises en compte. Les inscriptions consulaires ne sont pas prises en compte. 2) Parmi les différentes inscriptions successives décrites, l'indicatrice d'état (actif ou non) permet de repérer la situation à la date de production de la base d'études, donc en fin de période. 3) Les dates de fin d'inscription n'étaient pas mentionnées dans les premières bases d'études « EDP+ ». Or une nouvelle inscription peut faire suite à un déménagement, sans qu'il y ait eu d'interruption d'inscription ou faire suite à une interruption d'inscription suite à une radiation. Il est donc nécessaire de disposer des dates de fin d'inscription pour déterminer quels individus « EDP » sont inscrits une année donnée (sauf pour la dernière année d'observation). L'information devrait être intégrée dans les prochaines bases d'études.	

Table : RP68	**Unité observée :** L'individu « EDP »
Source : ***Recensement de* 1968** (bulletin individuel de l'individu « EDP »)	
Caractéristiques de l'individu « EDP » : lieu de résidence, sexe, catégorie de population au recensement, situation matrimoniale, lien avec la "personne de référence du ménage", année de naissance et lieu de naissance, nationalité, lieu de résidence au recensement de 1962, niveau de diplôme, âge de fin d'études, type d'activité, statut, catégorie socioprofessionnelle, lieu de travail, secteur d'activité, année d'installation en France pour les immigrés. ***Variables géographiques complémentaires :*** tranche d'unité urbaine de la commune de résidence (selon la géographie et la population communale de 1968). ***<u>Remarques</u> :*** 1) La codification et la saisie des bulletins ont été réalisées spécifiquement pour l'EDP. 2) La catégorie socioprofessionnelle a été codée selon les modalités de la nomenclature des PCS de 1982.	

Table : RP75	Unité observée : L'individu « EDP »

Source : ***Recensement de*** **1975** (feuille de logement, bulletins individuels de l'individu « EDP » et des autres membres du ménage)

Caractéristiques de l'individu « EDP » *:* information analogue à la table du recensement de 1968 + type de migration par rapport au recensement antérieur (même logement, autre logement dans même commune, …).

Caractéristiques de son logement *:* statut d'occupation du logement, nombre de pièces et niveau de confort ; type d'immeuble et nombre de logements, année d'achèvement.

Caractéristiques de son ménage *:* structure du ménage, nombre de personnes.

Caractéristiques de sa famille *:*
- structure de la famille ;
- caractéristiques du **« chef de famille »** : sexe, âge, nationalité, état matrimonial, diplôme, statut d'emploi et catégorie socioprofessionnelle ;
- caractéristiques du **« conjoint » de l'individu « EDP »** s'il est cohabitant : sexe, âge, nationalité, état matrimonial, diplôme, type d'activité, statut d'emploi et catégorie socioprofessionnelle, lieu de naissance et résidence antérieure.

Variables géographiques complémentaires : tranche d'unité urbaine d'appartenance de la commune de résidence et de la commune de résidence antérieure (selon la géographie et la population communale de 1975).

Remarques :

1) Certaines informations résultent d'une codification et d'une saisie réalisées spécifiquement pour l'EDP (variables « brutes » suffixées « 75 ») ; d'autres, parfois redondantes, des traitements réalisés pour l'exploitation du recensement (variables « redressées, suffixées « _75 »).

2) Une variable permet de repérer les personnes recensées pour lesquelles le rapprochement avec les informations « redressées » n'a pas été possible (variable NRP75).

3) Une variable de pondération permet d'assurer la représentativité des tabulations faites sur les informations « redressées » en tenant compte du fait que les exploitations du recensement n'ont été faites que sur un échantillon dans certaines communes (variable SOND_75).

4) Deux variables ont été construites spécifiquement pour l'EDP pour repérer les changements de profession et d'établissement employeur entre les recensements de 1968 et de 1975.

5) Les informations retenues pour décrire le conjoint sont plus nombreuses que celles retenues pour décrire le chef de famille.

6) Le « chef de famille » est la personne EDP dans 25% des cas, son conjoint dans 26% des cas et un de ses parents dans 34% des cas. La personne « EDP » est « hors famille » dans 16% des cas.

7) La profession déclarée dans le « code des métiers » de 1975 est disponible pour les personnes relevant de l'échantillon au cinquième constitué pour les premières exploitations nationales du recensement. La variable EB75 permet de repérer les individus concernés.

8) Certaines variables géographiques sont codées à partir du code officiel géographique de 1975 pour les individus relevant de l'échantillon au cinquième et du code géographique de 1982 pour les autres.

9) La catégorie socioprofessionnelle de l'individu « EDP » a été codée selon les modalités de la nomenclature des PCS de 1982. La catégorie socioprofessionnelle du conjoint et du chef de famille sont codées dans la nomenclature des CSP de 1969 (voir « Le code des catégories socioprofessionnelles », *Economie et Statistique*, n°4, septembre 1969).

10) Si la famille est composée d'un couple avec ou sans enfant, le chef de famille est l'époux. Si la famille est composée d'un seul adulte avec enfants, le chef de famille est l'adulte.

Table : RP82	**Unité observée :** L’individu « EDP »
Source : ***Recensement de* 1982** (bulletin individuel de l’individu « EDP » et, pour un quart des recensés, feuille de logement et bulletins individuels des autres membres du ménage)	

Caractéristiques de l’individu « EDP » : informations analogues à la table du recensement de 1975 mais suppression de l’année d’installation en France pour les immigrés (du fait de l’évolution du bulletin du recensement).

Caractéristiques de son logement *:* informations analogues à la table du recensement de 1975.

Caractéristiques de son ménage *:* informations analogues à la table du recensement de 1975.

Caractéristiques de sa famille *:*
- structure de la famille ;
- caractéristiques du **« père de famille »** et de la **« mère de famille »:** sexe, âge, nationalité, état matrimonial, situation d’activité, statut d’emploi et catégorie socioprofessionnelle ;

Variables géographiques complémentaires : tranche d'unité urbaine d’appartenance de la commune de résidence et de la commune de résidence antérieure (selon la géographie et la population communale de 1982).

Remarques :

1) Certaines informations résultent d’une codification et d’une saisie réalisées spécifiquement pour l’EDP (variables « brutes » suffixées « 82 ») ; d’autres, parfois redondantes, des traitements réalisés pour l’exploitation du recensement (variables « redressées, suffixées « _82 »).

2) Les variables « redressées », donc notamment les informations sur le niveau « logement », « ménage » et « famille », ainsi que le diplôme au niveau « individu », ne sont disponibles que pour un quart des personnes recensées (sondage au quart). Une variable permet de repérer les personnes concernées appartenant à cet échantillon (ECH82) ; une autre permet de repérer les personnes concernées pour lesquelles le rapprochement n’a pas été possible (variable NRP82).

3) Deux variables ont été construites spécifiquement pour l’EDP pour repérer les changements de profession et d’établissement employeur entre les recensements de 1975 et de 1982.

4) La catégorie socioprofessionnelle a été codée d’après la nomenclature des PCS de 1982.

5) Si la famille est composée d’un couple avec ou sans enfant, le chef de famille est l’époux. Si la famille est composée d’un seul adulte avec enfants, le chef de famille est l’adulte.

Table : RP90	**Unité observée :** L'individu « EDP »

Source : *Recensement de* 1990 (feuille de logement, bulletins individuels de l'individu « EDP » et des autres membres du ménage)

Caractéristiques de l'individu « EDP » : informations analogues à la table du recensement de 1982 mais suppression de l'âge de fin d'études et information plus agrégée sur le diplôme. Intégration d'informations utilisées pour coder la PCS (profession, nature de l'établissement, position professionnelle déclarée), ainsi que du temps de travail, des condition d'emploi, du mode de cohabitation.

Caractéristiques de son logement *:* informations analogues à la table du recensement de 1982.

Caractéristiques de son ménage *:* informations analogues à la table du recensement de 1982, avec ajout du type de ménages (combinaison de la structure simplifiée du ménage et du nombre de personnes).

Caractéristiques de sa famille *:*
- structure de la famille ; nombre d'enfants cohabitant présumés de l'individu « EDP » ;
- caractéristiques du **« père de famille »** et de la **« mère de famille »** : informations analogues à la table du recensement de 1982 ainsi que les conditions d'emploi, le secteur d'activité, le diplôme, le lieu de naissance, le lieu de résidence antérieure ;
- caractéristiques de la **« mère présumée » de l'individu « EDP »** quand elle cohabite : indicatrice de présence, année de naissance et lieu de naissance ;
- caractéristiques des **« enfants » cohabitants présumés de l'individu « EDP »** : sexe, date de naissance, lieu de naissance.

Variables géographiques complémentaires : tranche d'unité urbaine d'appartenance de la commune de résidence et de la commune de résidence antérieure ; zone d'emploi du lieu de résidence et du lieu de travail ; indicateur rural/urbain pour la commune de résidence ; type de commune dans la typologie en aire urbaine (codification selon la géographie, la population communale et les navettes domicile-travail de 1990) ; indicatrice de résidence dans une zone urbaine sensible (selon le zonage de 1990) et typologie « Tabard » des quartiers.

<u>Remarques :</u>

1) Certaines informations correspondent aux variables de saisie des chaînes d'exploitation du recensement (variables « brutes » suffixées « 90 ») ; d'autres, parfois redondantes, sont issues de l'exploitation « légère » et « lourde » du recensement et incluent donc des redressements et des imputations en cas de non-réponse (variables « redressées », suffixées « _90 »).

2) 48 500 individus « EDP » sont repérés comme résidant en ZUS à la date du recensement.

3) L'individu « EDP » est le « père de famille » dans un quart des cas ; il est la « mère de famille » dans un peu plus d'un quart des cas ; il est enfant de la famille dans un tiers des cas, le solde concernant les individus « hors famille ». Un tiers des individus « EDP » recensés cohabite avec au moins un de ses enfants présumés.

3) Deux variables ont été construites spécifiquement pour l'EDP pour repérer les changements de profession et d'établissement employeur entre les recensements de 1982 et de 1990.

4) Un « ménage » au sens du recensement regroupe toutes les personnes qui habitent habituellement dans le logement. Un ménage complexe peut ainsi regrouper plusieurs « familles » ou une « famille » et des personnes isolées.

Table : RP99	**Unité observée :** L’individu « EDP »

Source : *Recensement de* 1999 (feuille de logement, bulletins individuels de l’individu « EDP » et des autres membres du ménage)

Caractéristiques de l’individu « EDP » *:* informations analogues à la table du recensement de 1990, avec ajout de l’année d’arrivée en France, du niveau d’études, du lieu d’études pour les étudiants, et repérage des originaires des DOM.

Caractéristiques de son logement *:* informations analogues à la table du recensement de 1990, avec ajout de l’année d’emménagement et de la catégorie de communauté.

Caractéristiques de son ménage *:* informations analogues à la table du recensement de 1990, avec ajout des :
- caractéristiques de la **« personne de référence du ménage »** : sexe, lieu de naissance, nationalité, lieu de résidence antérieur, indicateur de qualité d’immigré, année d’arrivée en métropole, état matrimonial, diplôme, niveau d’études, situation d’activité, statut d’emploi, lieu de travail, conditions d’emploi, catégorie socioprofessionnelle, profession, secteur d’activité (à noter : l’âge et l’année de naissance ne sont pas disponibles)
- caractéristiques du **« conjoint de la personne de référence »** : mêmes informations que pour la personne de référence.

Caractéristiques de sa famille *:*
- structure de la famille ; nombre d’enfants cohabitants présumés de l’individu « EDP » ;
- caractéristiques du **« père de famille »** et de la **« mère de famille »** : informations analogues à la table du recensement de 1990 mais suppression de l’âge ;
- caractéristiques de la **« personne de référence de la famille »** : catégorie socioprofessionnelle, situation d’activité, statut d’emploi, diplôme, niveau d’études, lieu de naissance, lieu de résidence antérieure
- caractéristiques de la **« mère présumée » de l’individu « EDP »** quand elle cohabite : informations analogues à la table du recensement de 1990 ;
- caractéristiques des **« enfants » cohabitants présumés de l’individu « EDP »** : informations analogues à la table du recensement de 1990.

Variables géographiques complémentaires : tranche d'unité urbaine d’appartenance de la commune de résidence, de la commune d’études, de la commune du lieu de travail, indicateur rural/urbain pour la commune de résidence ; type de commune dans la typologie en aire urbaine (codification selon la géographie, la population communale et les navettes domicile-travail de 1999) ; indicatrice de résidence dans une zone urbaine sensible (selon le zonage de 1999) et typologie « Tabard » des quartiers.

<u>Remarques :</u>

1) Certaines informations correspondent aux variables de saisie des chaînes d’exploitation du recensement (variables « brutes » suffixées « 99 ») ; d’autres, parfois redondantes, sont issues de l’exploitation « légère » et « lourde » du recensement et incluent donc des redressements et des imputations en cas de non réponse (variables « redressées », suffixées « _99 »).

2) L’individu « EDP » est « personne de référence du ménage » dans 41% des cas, son conjoint dans 25% des cas et son enfant dans 32% des cas.

3) L’individu « EDP » est « personne de référence de la famille » dans 28% des cas, son conjoint dans 25% des cas et son enfant dans 32% des cas. Dans 15% des cas, il est « hors famille ».

4) 40 500 individus « EDP » sont repérés comme résidant en ZUS à la date du recensement.

5) Un « ménage » au sens du recensement regroupe toutes les personnes qui habitent habituellement dans le logement. Un ménage complexe peut ainsi regrouper plusieurs « familles » ou une « famille » et des personnes isolées.

Table : **EARaaaa_INDIVIDU** (aaaa=2004, …)	**Unité observée :** Les individus des « ménages » où réside au moins un individu « EDP »

Source : ***Enquête annuelle de recensement de l'année aaaa*** (bulletins individuels de l'individu « EDP » et des autres membres du ménage)

Caractéristiques des individus du ménage *:* sexe, année et lieu de naissance, indicateur de vie en couple, diplôme, nationalité, état matrimonial, lieu de résidence, profession, situation d'activité, catégorie socioprofessionnelle, résidence antérieure, année d'arrivée en France.

Variables géographiques complémentaires : code IRIS (notamment pour les « grandes » communes de métropole), code ILOT (« grandes » communes des départements d'outre-mer) et code District (« petites » communes). Type de commune.

Remarques :

1) Désormais les informations collectées sont homogènes sur toutes les personnes du ménage. La structure retenue doit permettre de réduire les évolutions de contenu d'un recensement à l'autre et éviter les redondances possibles qui existaient quand l'individu « EDP » était personne de référence de la famille ou du ménage, ou son conjoint. Pour l'utilisateur, la récupération des informations des personnes liées demande plus de manipulation. Une indicatrice permet de repérer les individus « EDP ». Plusieurs variables permettent de préciser la position de la personne observée au sein du ménage et des familles.

2) Certaines informations correspondent aux variables de saisie des chaînes d'exploitation du recensement (variables « brutes » suffixées « _X ») ; d'autres, parfois redondantes, sont issues de l'exploitation « principale » et « complémentaire » du recensement et incluent donc des redressements et des imputations en cas de non réponse (variables « redressées »).

3) Le numéro d'identifiant du logement et le numéro d'identifiant de la famille permettent de relier les personnes aux observations qui les concernent dans la table FAMILLE et dans la table LOGEMENT.

4) Deux variables de pondération de l'enquête annuelle sont proposées : le poids de l'exploitation principale et le poids de l'exploitation complémentaire. Les autres variables de pondération disponibles dans les fichiers du recensement ainsi que les variables utiles pour la stratification de la collecte n'ont pas été intégrées dans l'EDP (variable distinguant les personnes recensées dans les « petites » ou les « grandes » communes, les « petites » ou les « grandes » adresse, poids de tirage dans les groupes de rotation, poids de cumul, poids pour les restitutions aux communes).

Table : **EARaaaa_FAMILLE** (aaaa=2004, …)	**Unité observée :** Familles des « ménages » où réside au moins un individu « EDP »

Source : ***Enquête annuelle de recensement de l'année aaaa*** (variables calculées)

Caractéristiques de la famille *:* nombre de personnes, nombre d'enfants de la famille, nombre de personnes actives, nombre de personnes immigrées, type de famille.

Variables géographiques complémentaires : type de commune.

Remarque :

1) Le numéro d'identifiant du logement et le numéro d'identifiant de la famille permettent de relier les familles décrites aux observations qui les concernent dans la table INDIVIDU et dans la table LOGEMENT.

2) Un « ménage » au sens du recensement regroupe toutes les personnes qui habitent habituellement dans le logement. Un ménage complexe peut ainsi regrouper plusieurs « familles » ou une « famille » et des personnes isolées.

Table : EARaaaa_LOGEMENT (aaaa=2004, …)	**Unité observée :** Logement où réside au moins un individu « EDP »
Source : ***Enquête annuelle de recensement de l'année aaaa*** (feuille de logement, bulletin individuel de la personne de référence du ménage)	

Caractéristique du logement *:* type de logement, date d'achèvement, appartenance à un organisme HLM, nombre de pièces, moyen de chauffage.

Caractéristique du ménage *:* type de ménages, nombre de personnes, statut d'occupation, date d'emménagement, structure familiale du ménage.

Caractéristique de la « personne de référence du ménage » *:* sexe, âge, nationalité, pays de naissance, indicateur de qualité d'immigré, indicateur de vie en couple, diplôme, situation matrimoniale, situation d'activité, travail à temps partiel.

Variables géographiques complémentaires : tranche d'unité urbaine d'appartenance de la commune de résidence (géographie et population de 1999), type de commune.

<u>Remarques :</u>

1) Certaines informations correspondent aux variables de saisie des chaînes d'exploitation du recensement (variables « brutes » suffixées « _X ») ; d'autres, parfois redondantes, sont issues de l'exploitation « principale » et « complémentaire » du recensement et incluent donc des redressements et des imputations en cas de non-réponse (variables « redressées »).

2) Un numéro d'identifiant du logement permet de relier les logements aux familles et aux individus décrits dans les tables FAMILLE et INDIVIDU.

3) La variable de pondération de l'exploitation principale de l'enquête annuelle est proposée dans cette table. Les autres variables de pondération disponibles dans les fichiers du recensement ainsi que les variables utiles pour la stratification de la collecte n'ont pas été intégrées dans l'EDP (variable distinguant les personnes recensées dans les « petites » ou les « grandes » communes, les « petites » ou les « grandes » adresse, poids de tirage dans les groupes de rotation, poids de cumul, poids pour les restitutions aux communes).

ANNEXE 2 - Les bulletins d'état civil[78]

Depuis la Révolution, l'état civil est tenu par les communes, dans le cadre général fixé depuis Napoléon, par le code civil. Le titre II du Livre Ier est entièrement consacré aux actes d'état civil : les principes généraux (chapitre Ier), les actes de naissance (chapitre II), les actes de mariage (chapitre III), les actes de décès (chapitre IV), les rectifications de l'état civil (chapitre VII) et des cas particuliers concernant les militaires et les marins (chapitre V) ou de personnes nées à l'étranger acquérant la nationalité française (chapitre VI).

L'ensemble des dispositions législatives et réglementaires qui fixent tant les principes que les modes de gestion pratique de l'état civil, y compris la façon de tenir les registres et de délivrer les actes, est rassemblé dans une instruction générale relative à l'état civil du ministre de la justice. L'instruction générale du 21 septembre 1955, modifiée à de multiples reprises pour tenir compte de l'évolution règlementaire et technologique, précise dans la section 6 de son chapitre II, intitulé « bulletins statistiques », qu' « en vue de permettre la statistique générale de la France, les officiers de l'état civil remplissent, lors de la rédaction de chaque acte, un bulletin spécial comprenant, outre les énonciations principales de l'acte, certaines indications précisant notamment la situation de famille, le degré d'instruction et la catégorie professionnelle des intéressés ».

L'instruction générale du 21 septembre 1955 a été abrogée par une nouvelle instruction générale, datée du 11 mai 1999. Elle prévoit toujours l'envoi par les communes de bulletins statistiques à destination de l'INSEE pour les évènements d'état civil enregistrés dans la commune, avec comme double finalité, l'établissement de statistiques et la mise à jour du répertoire national d'identification des personnes physiques.

« Section 6 - Bulletins statistiques »

« En vue de permettre l'établissement des statistiques du mouvement de la population, la tenue et la mise à jour du Répertoire national d'identification des personnes physiques (R.N.I.P.P.) les officiers de l'état civil remplissent des bulletins statistiques de l'état civil contenant, outre les énonciations de l'acte, certains renseignements complémentaires portant notamment sur la situation familiale et l'activité professionnelle des intéressés. »
« Ces bulletins sont établis soit lors de l'enregistrement d'un acte, d'après les indications du déclarant, soit à l'occasion d'une transcription ou d'une mention en marge. »
« Il existe huit modèles de bulletin : [liste]
« Outre leur utilisation à des fins statistiques, certains bulletins servent à la gestion administrative. Les bulletins n°5 (naissances), n°4 (reconnaissance), n°1 (transcriptions), n°1 bis (mentions marginales), et n°7 bis (décès) sont utilisés par l'INSEE pour la mise à jour du Répertoire national d'identification des personnes physiques (décret n° 82-103 du 22 janvier 1982). »

« [modalités de transmission à l'INSEE] »
« Les bulletins de naissance (n°5) doivent être envoyés à l'INSEE le jour même de la rédaction de l'acte de naissance ; ceux de reconnaissance (n°4) et ceux de décès (n°7 bis) dans un délai maximum de huit jours ; les autres bulletins (n°1, 1 bis, 2 et 6) doivent être regroupés dans des envois mensuels et envoyés au plus tard cinq jours suivant la fin du mois. (Décret n°82-103 du 22 janvier 1992 modifié relatif au Répertoire national d'identification des personnes physiques et instituant des délais de transmission d'informations d'état civil). »
« Cette matière dépassant le cadre de la présente instruction, aucune référence ne sera faite à l'établissement des bulletins statistiques lors de l'examen des règles particulières aux divers actes de l'état civil. »

[78] Cette annexe a été rédigée avant la promulgation de la loi n° 2013-404 du 17 mai 2013 ouvrant le mariage aux couples de personnes de même sexe.

Les bulletins statistiques sont remplis par la commune, la plupart des informations demandées étant contenues dans les actes d'état civil. Ces derniers sont eux-mêmes rédigés sur la foi du ou des déclarants. Concernant l'identification des personnes, l'article 34 du code civil stipule notamment que :

> *« tout acte d'état civil doit énoncer l'année, le jour et l'heure où il est établi, les prénoms, noms, professions et domiciles de tous ceux qui y seront dénommés, ainsi que les dates et lieux de naissance :*
> *- des père et mère dans les actes de naissance et de reconnaissance ;*
> *- de l'enfant dans les actes de reconnaissance ;*
> *- des époux dans les actes de mariage ;*
> *- du décédé dans les actes de décès »*

Concernant la profession figurant dans l'acte d'état civil, l'instruction générale précise :

> *« La profession qui doit être indiquée s'entend non seulement de celle actuellement exercée par les intéressés, mais également de celle qu'ils exerçaient en dernier lieu avant de cesser leur activité professionnelle. Il convient, dans cette hypothèse, de préciser ancien, en retraite, ou honoraire, ou en cas de chômage, de préciser sans emploi. Lorsqu'une personne n'a jamais exercé une profession, il convient d'indiquer sans profession ».*

Les bulletins statistiques intégrés à l'échantillon démographiques ont fait l'objet de deux révisions d'ensemble :

- Une première, mise en œuvre en 1998 : elle simplifie les questions sur la profession et sur l'adresse, faisant disparaître certaines variables.

- Une seconde, mise en œuvre en septembre 2008 : elle harmonise les questions posées dans les formulaires proches (bulletin de jugement déclaratif de décès et bulletin de décès, d'une part ; bulletin de naissance, bulletin de jugement déclaratif de naissance et bulletin d'enfant sans vie, d'autre part) ; elle intègre les évolutions législatives qui ont mis fin à la distinction entre enfant naturel et enfant légitime et ont ouvert la possibilité du choix du nom de famille ; elle remplace le bulletin de jugement par trois modèles (jugement déclaratif de naissance, jugement déclaratif de décès, jugement d'adoption plénière) ; elle supprime le bulletin de reconnaissance.

Naissance (bulletin n°5)

Selon le code civil, les déclarations de naissance doivent être faites dans les trois jours suivant l'accouchement à l'officier de l'état civil de la commune de naissance. Si une naissance n'est pas déclarée dans le délai légal, l'inscription sur les registres nécessite un jugement préalable et une mention sommaire est alors inscrite en marge à la date de la naissance (article 55). C'est au père qu'il appartient de déclarer la naissance ou, à défaut, tout autre personne ayant assisté à l'accouchement (article 56).

L'acte de naissance énonce le jour, l'heure et le lieu de la naissance, le sexe de l'enfant, les prénoms donnés, le nom de famille[79], les prénoms, noms, âges, professions et domiciles des père et mère (article 57). Il comporte également les dates et lieux de naissance des père et mère (article 34).

[79] Depuis 2005, le 1er janvier 2005, le nom n'est plus forcément celui du père : « Lorsque la filiation d'un enfant est établie à l'égard de ses deux parents au plus tard le jour de la déclaration de sa naissance ou par la suite mais simultanément, ces derniers choisissent le nom de famille qui lui est dévolu : soit le nom du père, soit le nom de la mère, soit leurs deux noms accolés dans l'ordre choisi par eux dans la limite d'un nom de famille pour chacun d'eux. En l'absence de déclaration conjointe à l'officier de l'état civil mentionnant le choix du nom de l'enfant, celui-ci prend le nom de celui de ses parents à l'égard duquel sa filiation est établie en premier lieu et le

Jusqu'en mars 1993, seuls les enfants vivants au moment de la déclaration donnaient lieu à un acte de naissance. Dans le cas contraire on enregistrait un acte d'enfant sans vie, que l'enfant ait vécu ou non. Depuis mars 1993, l'officier de l'état civil enregistre un acte de naissance si l'enfant a respiré puis un acte de décès s'il est mort avant la déclaration de sa naissance. Si l'enfant n'est pas né vivant et viable, il enregistre un acte d'enfant sans vie. Depuis cette date, le bulletin n°5 s'intitule donc « bulletin de naissance » au lieu de « bulletin d'enfant déclaré vivant ».

Un acte tenant lieu d'acte de naissance est dressé pour toute personne née à l'étranger qui acquiert ou recouvre la nationalité française à moins que l'acte dressé à sa naissance n'ait déjà été porté sur un registre conservé par une autorité française. Cet acte énonce les nom, prénoms et sexe de l'intéressé et indique le lieu et la date de sa naissance, sa filiation, sa résidence à la date de l'acquisition de la nationalité française (article 98).

Le bulletin statistique n°5 est utilisé pour transmettre les informations de l'acte de naissance à l'INSEE. Il sert également à procéder à l'inscription de la personne au répertoire national d'identification des personnes physiques. Il comprend des renseignements sur :

- L'enfant : nom, prénom, sexe, date de naissance, déclarations conjointe de choix du nom (si oui, date de déclaration commune du choix) ;
- Chacun des parents : nom, prénom, date et lieu de naissance, activité, adresse, nationalité ;
- La filiation : mariage des parents (date et lieu) ; reconnaissance par le père (date), la mère (date) ou conjointe (date), accouchement anonyme ou enfant trouvé ;
- Les conditions de l'accouchement (nombre d'enfants issus de l'accouchement, établissement spécialisé, ailleur avec assistance médical, ailleurs sans assistance) ;
- Le nombre d'enfant antérieur de la mère et date de naissance du dernier enfant.

Le bulletin n°5 a été visé par le CNIS (n°201 PA 002 EC, valable de 2011 à 2015). Il a un double usage : l'élaboration de statistiques et à la mise à jour du RNIPP.

Jusqu'en 1997, le bulletin statistique précisait la commune de mariage des parents (quand ils étaient mariés).

Jusqu'en 2008, le bulletin statistique proposait la qualité juridique de l'enfant : « légitime », « naturel non reconnu », « naturel reconnu par le père seul », « naturel reconnu par la mère seul », « naturel reconnu par le père et la mère ».

Depuis 2008, le bulletin statistique demande l'adresse du père, comme il le faisait pour la mère jusqu'alors.

Enfants sans vie (bulletin n°6)

Jusqu'en 1919, l'enfant nouveau né devait être présenté à l'officier d'état civil. Si celui-ci était décédé, il dressait alors un « acte de présentation d'un enfant sans vie » : « lorsque le cadavre d'un enfant dont la naissance n'a pas été enregistrée sera présenté à l'officier de l'état civil, cet officier n'exprimera pas qu'un tel enfant est décédé, mais seulement qui lui a été présenté sans vie. Il recevra de plus la déclaration des témoins touchant les noms, prénoms, qualités et demeure des père et mère de l'enfant, et la désignation des an, jour et heure auquel l'enfant est sorti du sein de sa mère » (décret du 4 juillet 1806). Pour les enfants mort-nés ou nés vivants mais décédés avant leur présentation à l'officier d'état civil, il n'y avait donc ni acte de naissance, ni acte de décès. En 1919, l'obligation de présenter l'enfant est supprimée et l'acte devient un « acte d'enfant sans vie », mais les principes généraux ne changent pas.

Depuis la loi du 8 janvier 1993, applicable depuis mars 1993, l'article 79-1 du Code civil a restreint la rédaction de l'acte d'enfant sans vie aux seuls enfants dont il était établi qu'ils n'étaient pas nés vivant et viable. Dans le cas contraire, « lorsqu'un enfant est décédé avant que sa naissance ait été déclarée à l'état civil, l'officier de l'état civil établit un acte de naissance et un acte de décès sur production d'un

---

nom de son père si sa filiation est établie simultanément à l'égard de l'un et de l'autre […]». Article 311-21 du code civil.

certificat médical indiquant que l'enfant est né vivant et viable et précisant les jours et heures de sa naissance et de son décès ». En l’absence de certificat médical précisant la viabilité de l’enfant, comme auparavant, aucun acte de naissance n’est établit mais un acte d’enfant sans vie est dressé. Modifié dans son contenu, « il énonce les jour, heure et lieu de l'accouchement, les prénoms et noms, dates et lieux de naissance, professions et domiciles des père et mère et, s'il y a lieu, ceux du déclarant. L'acte dressé ne préjuge pas de savoir si l'enfant a vécu ou non ». L'enregistrement s'effectue dans la commune de naissance ou dans la commune où le corps de l'enfant se trouvait au moment où l'on a constaté qu'il était sans vie.

L’évolution de 1993 a conduit à préciser la notion de viabilité pour éviter la déclaration de fœtus qui auraient pu présenter quelques signes de vie. Dans un premier temps, des textes réglementaires ont considéré qu’une durée minimale de 180 jours de gestation ou 28 semaines de grossesse était nécessaire. Une circulaire du 30 novembre 2001 a modifié ces critères pour s’aligner sur ceux de l’Organisation mondiale de la santé : au moins 22 semaines de grossesse ou si le fœtus atteint un poids de 500 grammes. A défaut, il ne pouvait y avoir d’établissement d’acte d’enfant sans vie. En février 2008, la Cour de cassation a jugé ces critères plus restrictifs que le texte de la loi, laquelle ne mentionne aucun critère de poids ou de durée de la grossesse. Un décret du 20 août 2008 a donc mis fin à l’application de ces critères : les actes d’enfants sans vie peuvent désormais être établis à la demande des parents, en l’absence de certificat médical indiquant que l'enfant est né vivant et viable, quel que soit le poids du fœtus et la durée minimale de grossesse, sur la foi d’un certificat médical constatant l'existence d'un accouchement.

Cette fluctuation dans les définitions a eu des effets statistiques : le nombre d'enfants sans vie est passé de 4 000 en 2001 à près de 6 700 en 2002, soit un accroissement de plus de 67 %. Il a augmenté à nouveau de 14% en 2008 et 13% en 2007.

Le bulletin statistique n°6 est utilisé pour transmettre l’information à l’INSEE. Il comprend des renseignements sur :

- L’enfant : prénom, sexe ;
- L’accouchement : date de l’accouchement, nombre d’enfants issus de l’accouchement, condition de l’accouchement (établissement spécialisé, ailleurs avec assistance médicale, ailleurs sans assistance) ;
- Chacun des parents : nom, prénom, date et lieu de naissance, activité, adresse, nationalité ;
- La filiation : mariage des parents (date et lieu), accouchement anonyme ou enfant trouvé ;
- Le nombre d’enfant antérieur de la mère et date de naissance du dernier.

Le bulletin n°6 a été visé par le CNIS (n°201 PA 003 EC, valable de 2011 à 2015). Il n’a qu’un usage statistique.

Les informations intégrées dans l’échantillon démographique permanent dans version « BRPP1 » se limitent à la date de l’accouchement.

Mariages (bulletin n°2)

A l’occasion de tout mariage célébré en France, un acte de mariage est dressé. Il est rédigé dans la commune où le mariage est célébré, le jour même du mariage (« sur le champ », selon l’article 75 du code civil). Il énonce notamment : les prénoms, noms, professions, âges, dates et lieux de naissance, domiciles et résidences des époux ; les prénoms, noms, professions et domiciles des pères et mères ; le cas échéant, les prénoms et nom du précédent conjoint de chacun des époux (article 76 du code civil). La célébration du mariage et le nom du conjoint sont également mentionnés en marge de l'acte de naissance de chaque époux.

Le bulletin statistique n°2 est utilisé pour remonter l’information à l’INSEE. Il comprend des renseignements sur :

- Chacun des époux : nom, prénom, date et lieu de naissance, nationalité, état matrimonial antérieur (célibataire ; veuf ou divorcé avec la date de l’événement dans ces deux derniers cas) ;
- Le lieu de résidence probable (commune/pays) ;
- L’existence d’enfants en commun (si oui, combien).

Le bulletin a été visé par le CNIS (n°201 PA 001 EC, valable de 2011 à 2015). Il a un double usage : l'élaboration de statistiques et la mise à jour du RNIPP (pour les cas de changements de noms).

Jusqu'en 1997, le bulletin statistique de mariage contenait également la catégorie socioprofessionnelle de l'époux et de son père, la catégorie socioprofessionnelle de l'épouse et de son père ; les domiciles antérieurs de l'époux et de l'épouse.

Les bulletins de 2008 suppriment la liste des enfants légitimés par le mariage, la notion ayant disparu.

Reconnaissances (bulletin n°4)

Le code civil prévoit trois modalités d'établissement de la filiation : l'effet de la loi, la reconnaissance volontaire et la possession d'état constatée par un acte notarié. Elle peut aussi être établie par jugement, à la demande de l'enfant (par exemple, dans les cas de recherche de paternité quand le père présumé refuse la reconnaissance de l'enfant).

Depuis l'ordonnance du 4 juillet 2005, entrée en vigueur le 1er juillet 2006, la filiation de l'enfant avec sa mère est établie automatiquement par sa désignation dans l'acte de naissance de l'enfant. Elle n'a pas besoin de procéder à des démarches particulières. Sauf cas particuliers, si la mère est mariée, son mari est automatiquement présumé père de l'enfant et mentionné comme tel dans l'acte de naissance. Pour les couples non mariés, la filiation n'est pas établie automatiquement entre le père et l'enfant. Pour lui, une façon de l'établir est donc de reconnaître l'enfant.

Avant la réforme de 2005, la loi distinguait les enfants « légitimes » des enfants « naturels », qui n'avaient pas les mêmes droits, notamment en matière successorale. La filiation ne s'établissait pas de la même façon selon que les parents étaient mariés ou non. Sauf cas particuliers, les enfants nés de parents mariés étaient dits « légitimes ». L'indication du nom des parents mariés dans l'acte de naissance suffisait à prouver la filiation. En revanche, la filiation n'était pas établie automatiquement pour les enfants « naturels », nés hors mariage. En particulier, l'acte de naissance portant l'indication de la mère ne valait reconnaissance que si elle était corroborée par la possession d'état. La façon la plus simple d'établir la filiation était donc, tant pour la mère que pour le père, de procéder à une reconnaissance volontaire.

Avant la réforme de 2005, les enfants nés d'un couple non marié pouvaient toutefois être légitimés de plein droit par le mariage de leurs parents, à condition que la filiation ait déjà été établie avec chacun des deux parents. Dans le cas contraire, la légitimation nécessitait une reconnaissance de l'enfant dans un acte séparé, le jour du mariage (si la filiation n'était établit qu'après le mariage, la légitimation restait possible mais supposait un jugement préalable constatant la possession d'état d'enfant commun).

Avant la naissance, la reconnaissance peut s'effectuer dans n'importe quelle mairie. Dans ce cas, elle est mentionnée dans l'acte de naissance de l'enfant. A la naissance, la reconnaissance peut s'effectuer en même temps que la déclaration de naissance. Elle est alors contenue dans l'acte de naissance de l'enfant et ne donne pas lieu à d'acte complémentaire. Après la naissance, la reconnaissance peut s'effectuer dans n'importe quelle mairie. Elle est alors portée en marge de l'acte de naissance dans les registres de la commune de naissance de l'enfant et donne donc lieu à un bulletin de mention en marge.

Reconnaître son enfant : une démarche de plus en plus fréquente et de plus en plus souvent anticipée
Pascal Germé, Lucile Richet-Mastain, division Enquêtes et études démographiques,
Insee Première *N°1105 - octobre 2006*

« Comme dans tous les pays d'Europe occidentale, la hausse des naissances hors mariage correspond à la dissociation entre la procréation et le mariage intervenue dans la seconde moitié du xx[e] siècle. En France métropolitaine, 46,4 % des naissances étaient le fait de couples non mariés en 2004, contre seulement 8,5 % en 1974. Dans le même temps, reconnaître son enfant est devenu un acte de plus en plus fréquent.
[…]
Un enfant peut être reconnu pendant la grossesse, à sa naissance ou après. Entre 1974 et 2004, le nombre de reconnaissances prénatales est passé de 6 900 à 211 800 […]. Sur la même période, les reconnaissances après la naissance ont augmenté de 81 800 à 152 500 […]. Au total, sur les 379 300 enfants nés en France en 2004 de parents non mariés, plus de la moitié (59 %) ont été reconnus avant ou au moment de leur naissance par leurs deux parents et sont donc dans une situation proche sur le plan juridique de celle des enfants nés d'un couple marié. C'était le cas de 52 % des enfants nés hors mariage en 1999.

La reconnaissance peut être effectuée par le père seul, la mère seule ou les deux parents, conjointement ou séparément. Un peu plus de la moitié (51 %) des enfants nés hors mariage en 2004 ont été reconnus conjointement par leurs deux parents avant leur naissance (soit 87 % des naissances reconnues avant ou au moment de la naissance par les deux parents). Cette tendance s'est développée rapidement depuis le début des années quatre-vingt : c'était le cas de 8 % des enfants nés hors mariage en 1980, de 23 % en 1990 et de 35 % en 1994. »

Selon l'article 62 du code civil, l'acte de reconnaissance énonce les prénoms, nom, date de naissance ou, à défaut, âge, lieu de naissance et domicile de l'auteur de la reconnaissance. Il indique les date et lieu de naissance, le sexe et les prénoms de l'enfant ou, à défaut, tous renseignements utiles sur la naissance. Il est inscrit à sa date sur les registres de l'état civil. Les informations sur l'auteur de la reconnaissance sont portées en marge de l'acte de naissance de l'enfant.

Le bulletin statistique n°4 était utilisé pour remonter l'information à l'INSEE. Il comprenait des renseignements sur :

- La reconnaissance : date, circonstance (avant la naissance de l'enfant, pendant sa vie ou après son décès), les auteurs (père seul, mère seul, les deux) ;
- Chacun des parents : Nom, prénom, date et lieu de naissance, nationalité et adresse ;
- L'enfant : prénom, sexe, date et lieu de naissance ;
- L'existence d'une déclaration de changement de nom en même temps que la reconnaissance.

A la suite d'un changement de traitement informatique des bulletins de reconnaissance en 1993, les fichiers de l'état civil sont incomplets pour les années 1993 à 1995.

Depuis 2008, il n'y a plus de bulletin statistique de reconnaissance.

Décès (bulletin n°7)

Selon le code civil, « l'acte de décès est dressé par l'officier de l'état civil de la commune où le décès a eu lieu, sur la déclaration d'un parent du défunt ou sur celle d'une personne possédant sur son état civil les renseignements les plus exacts et les plus complets qu'il sera possible » (article 78). Le décès doit être déclaré dans les 24 heures suivant la constatation du décès. Au préalable, le décès doit donc être constaté par un médecin qui établit alors, si la cause de la mort n'est pas violente, un certificat de décès.

L'acte de décès énonce notamment le jour, l'heure et le lieu de décès ; les prénoms, nom, date et lieu de naissance, profession et domicile de la personne décédée ; les prénoms, noms, professions et domiciles de ses père et mère ; les prénoms et nom de l'autre époux, si la personne décédée était mariée, veuve ou divorcée (ou du partenaire en cas de pacte civil de solidarité depuis mai 2011). Il est fait mention du décès en marge de l'acte de naissance de la personne décédée (article 79).

Le décès donne lieu à la rédaction de deux bulletins statistiques : le bulletin n°7, destiné à l'Institut national de la santé et de la recherche médicale, pour le suivi des causes de décès, et le bulletin n°7 bis, destiné à l'Insee, pour les statistiques démographiques et la mise à jour du répertoire des personnes physiques. Le bulletin n°7bis contient donc l'identité du défunt alors que le bulletin n°7 est anonyme pour respecter le secret médical.

En pratique, le médecin qui constate le décès établit un certificat de décès qui inclut un volet destiné à préciser la ou les causes du décès, selon des règles de classification fixées par l'Inserm, en conformité avec les normes internationales. Ce volet confidentiel est cacheté par le médecin[80]. Le certificat médical est remis à l'officier d'état civil qui, après avoir dressé l'acte de décès, remplit les bulletins statistiques 7 et 7 bis. Le bulletin n°7 est ensuite transmis à l'autorité régionale de santé (auparavant, la direction départementale de l'Action sanitaire et sociale), avec le volet du certificat de décès indiquant les causes du décès. L'information est ensuite centralisée par l'Inserm. De son côté, le bulletin n°7bis est transmis directement à l'Insee.

Le bulletin statistique n°7 bis comprend des informations sur :

- Le défunt : nom, prénom, sexe, date et lieu de naissance, activité, nationalité, adresse, état matrimonial lors du décès (célibataire, marié, veuf, divorcé) ;
- Le décès : date du décès, lieu (logement, hôpital, clinique, maison de retraite, voie publique).

Il a été visé par le CNIS (n°201 PA 005 EC, valable de 2011 à 2015). Il sert à l'élaboration de statistique et à la mise à jour du RNIPP.

Le bulletin n°7 comprend les mêmes information que le bulletin n°7 bis, à l'exception des nom et prénom. Il a lui aussi été visé par le CNIS (n°201 PA 004 EC, valable de 2011 à 2015).

Transcriptions et mentions en marge (bulletin n°1)

Certains évènements d'état civil enregistrés en dehors de la commune doivent être transcrits sur ses registres pour avoir leurs effets en France. Il en est de même dans le cas de jugements pris en cas de défaut de déclaration ou pour certains évènements faisant l'objet d'actes authentiques.

En particulier :

- Lorsque la déclaration de naissance n'a pas été faite dans les trois jours suivant l'accouchement, un jugement est nécessaire ; il est transcrit sur les registres de la commune de naissance (ou du service central d'état civil du ministère des affaires étrangères pour les évènements survenus à l'étranger).

- Lorsque le décès d'une personne ne peut être constaté, par exemple parce que le corps du défunt n'a pas été retrouvé, un jugement est nécessaire ; il est transcrit sur les registres de la commune de décès présumé ou, à défaut, de la dernière commune de résidence connue (ou du service central d'état civil du ministère des affaires étrangères pour les décès survenus à l'étranger).

- Les jugements d'adoption plénière sont transcrits sur les registres de la commune de naissance de l'adopté, dont l'acte de naissance initiale est alors annulé ; dans le cas d'adoption d'enfants étrangers par des parents français, la transcription est effectuée sur les registres du service central d'état civil du ministère des affaires étrangères.

- Les mariages effectués à l'étranger impliquant au moins un conjoint français doivent être transcrits pour être opposables à des tiers en France.

[80] Depuis juillet 2006, la certification des causes de décès peut aussi se faire de façon électronique via une interface dédiée mise en place par l'Inserm. La partie médicale anonyme est alors transmise de façon cryptée à l'Inserm. La partie administrative du certificat, sans la cause du décès, est imprimée à destination de l'officier d'état civil.

Certaines transcriptions conduisent à l'établissement d'un bulletin statistique transmis à l'Insee, à des fins statistiques et pour la mise à jour du RNIPP. Depuis septembre 2008, sont ainsi concernés :

- La transcription d'un jugement d'adoption plénière : bulletin statistique n°1a ;
- La transcription d'un jugement déclaratif de naissance : bulletin statistique n°1b ;
- La transcription d'un jugement déclaratif de décès ou d'absence : bulletin statistique n°1c.

Le bulletin statistique n°1a (jugement d'adoption plénière) contient des informations sur :

- La date de jugement,
- L'état civil de la personne avant l'adoption (identité, sexe, date et lieu de naissance),
- L'identité après l'adoption,
- L'état civil et la profession de chacun des parents adoptifs, leur adresse et leur statut matrimonial.

Le bulletin statistique n°1b (jugement déclaratif de naissance) contient des informations sur :

- La date de jugement,
- L'état civil de l'enfant (identité, sexe, date et lieu de naissance),
- L'état civil et la profession de chacun des parents, leur adresse et leur statut matrimonial.

Le bulletin statistique n°1c (jugement déclaratif de décès ou d'absence) contient des informations sur :

- La date de jugement,
- L'état civil de la personne, sa profession et son adresse,
- La date et le lieu du décès ou de la déclaration d'absence,

Les trois bulletins ont été visés par le CNIS (respectivement n°201 PA 007 EC, n°201 PA 008 EC, n°201 PA 009 EC, valables de 2011 à 2015).

Avant septembre 2008, le bulletin statistique de transcription était unique et servait pour les trois types de transcription.

Mentions en marge

Certains événements d'état civil doivent être mentionnés en marge d'autres actes, par exemple :

- La reconnaissance, en marge de l'acte de naissance de l'enfant reconnu (pour les reconnaissances établies devant un officier d'état civil ou pour les reconnaissances établies devant notaire et transmises à un officier d'état civil sur demande des intéressés) ;
- Le mariage, en marge de l'acte de naissance de chacun des époux ;
- Le divorce, en marge de l'acte de naissance et de l'acte de mariage des divorcés ;
- Les modifications de l'état civil, notamment le patronyme, en marge de l'acte de naissance ;
- Le décès, en marge de l'acte de naissance ;
- Les modifications de date de décès, en marge de l'acte de décès.

Certaines mentions conduisent à l'établissement d'un bulletin statistique transmis à l'Insee, à des fins statistiques et pour la mise à jour du RNIPP (notamment pour repérer si l'acte ou le jugement porté en mention a été déjà récupéré par l'Insee) : le bulletin statistique n°3. Ce bulletin précise que « l'Insee ne souhaite pas recevoir toutes les mentions apposées sur les actes de naissances ou de décès ».

Outre l'état civil initial de la personne concernée (identité, sexe, date et lieu de naissance), sont demandés :

- En cas de modification des éléments civils d'un acte de naissance : la date de la mention, la nature de l'élément modifié (nom de famille, prénoms, sexe ou date de naissance) et l'information modifiée.
- En cas de mention d'un mariage sur l'acte de naissance : la date et le lieu de l'événement.
- En cas de mention d'une reconnaissance sur l'acte de naissance : la date et le lieu de l'événement (pas d'information sur la personne qui effectue la reconnaissance).
- En cas de modification de la date de décès sur un acte de décès : la date de la mention, la date et le lieu du décès avant la mention et la date modifiée.
- En cas de mention annulant un acte de naissance ou de décès : la date de la mention, le numéro de l'acte annulé et la raison (notamment si l'annulation de l'acte de naissance fait suite à une adoption).

Le bulletin a été visé par le CNIS (n°201 PA 006 EC, valable de 2011 à 2015).

Avant septembre 2008, le bulletin de mention en marge couvrait dix catégories de mention.

Selon la documentation disponible, seule deux informations sont clairement identifiées comme issues de mentions en marge dans les bases d'études de l'EDP antérieures au projet « EDP+ » : la date de la mention et le fait qu'il s'agit ou non d'une mention de légitimation de l'enfant par mariage. Il se peut que les bulletins de mention en marge aient été utilisés pour alimenter ou actualiser des variables associées aux bulletins auxquelles la mention correspondait mais ce point n'a pu être éclairci.

Les évènements d'état civil de la base d'étude (jusqu'à 2006 inclus)

Type de bulletins	Nombre	Nombre maximum de cas décrits par individu « EDP »
Naissance d'une personne « EDP » bulletin n°5 « collecté » bulletin n°5 « fantôme » bulletin n°5 « confirmé » bulletin n°5 « erroné » bulletin n°1, jugement bulletin n°1, adoption	**299 929** 297 922 1 050 526 231 35 165	**1**
Mariage légitimant une personne « EDP » bulletin n°2 « collecté » bulletin n°2 « fantôme » bulletin n°2 « confirmé » bulletin n°2 « erroné »	**14 434** 13 745 360 168 161	**1**
Mariage d'une personne « EDP » bulletin n°2 « collecté » bulletin n°2 « fantôme » bulletin n°2 « confirmé » bulletin n°2 « erroné » **Nombre d'individus « EDP » concernés** avec 1 bulletin de mariage avec 2 bulletins avec 3 bulletins avec 4 bulletins	**209 228** 203 169 654 3 205 2 200 **196 795** 181 164 11 927 666 38	**4**
Naissance d'un enfant d'une personne « EDP » bulletin n°5 bulletin n°1, jugement bulletin n°1, adoption **Nombre d'individus « EDP » concernés** avec 1 bulletin de naissance avec 2 bulletins dont 3 bulletins dont 4 bulletins dont 5 bulletins …	**492 048** 491 854 48 146 **272 888** 122 628 100 384 36 915 9 127 2 481	**12**
Reconnaissance bulletin n°4 « collecté » bulletin n°4 « fantôme » bulletin n°4 « confirmé » bulletin n°4 « erroné » **Nombre d'individus « EDP » concernés** avec 1 bulletin de reconnaissance *reconnaissance de l'enfant « EDP »* *reconnaissance par un individu « EDP »* avec 2 bulletins *2 reconnaissance de l'enfant « EDP »* *2 reconnaissances par l'individu « EDP »* *cas mixte*	**100 488** 95 007 202 2 312 2 967 **80 920** 61352 *24 988* *36 364* 19 568 *3 990* *15 550* *28*	**2**
Naissance d'un enfant sans vie de d'une personne « EDP » bulletin n°6 « collecté » bulletin n°6 « fantôme » **Nombre d'individus « EDP » concernés** avec 1 bulletin d'enfants sans vie avec bulletins	**3 620** 3 615 5 **3 468** 3 316 152	**2**

Décès d'une personne « EDP »	**188 242**	**1**
bulletin n°7 bis « collecté »	87 050	
bulletin n°7 bis « fantôme » (repérage du décès au RNIPP)	98 892	
bulletin n°7 bis « confirmé »	1 108	
bulletin n°7 bis « erroné »	1 150	
bulletin n°1 « collecté »	41	
bulletin n°1 « fantôme » (repérage du décès au RNIPP)	1	
Mention en marge	**14 853**	**2**
mention légitimant l'enfant « EDP » par mariage des parents	2 145	
autre cas (non précisé)	12 708	
Nombre d'individus « EDP » concernés		
avec une mention en marge	14 436	
avec deux mentions en marges (ou plus)	14 019	
	417	

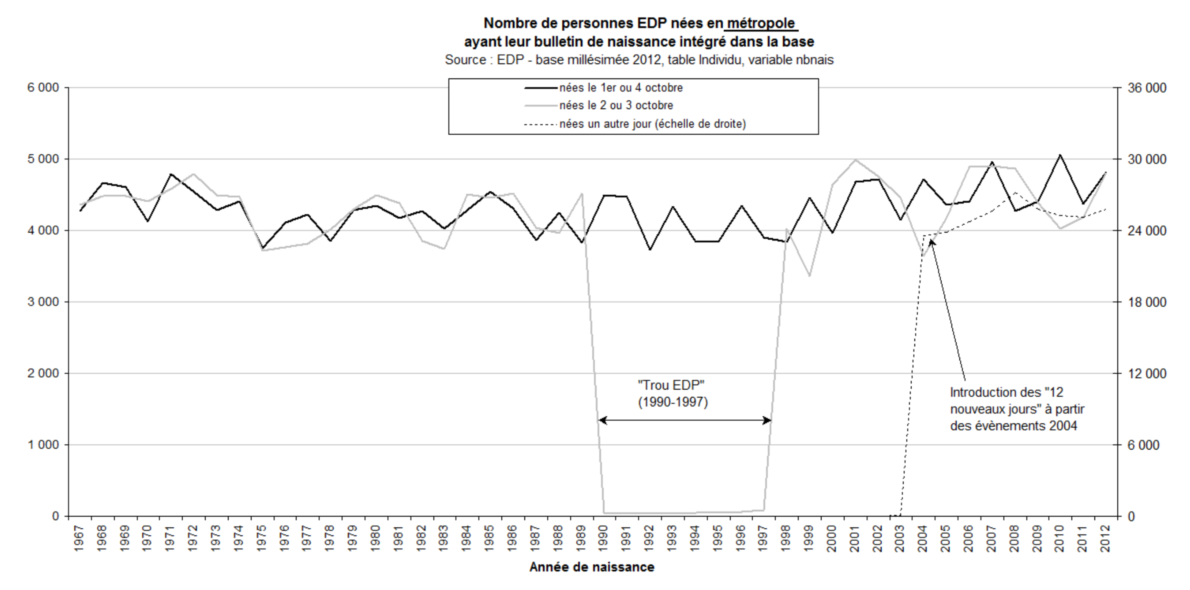
Notes :

- *Pour les naissances d'individus « EDP »* : on ne prend en compte ici qu'un bulletin de décès par personne (78 personnes ont deux bulletins et 1 personne a 3 bulletins) ; utilisation des variables STATNA et OBNA.
- *Pour les mariages d'individus « EDP »* : utilisation des variables STATNM(i) et NBM ; pour 6 individus l'EDP compte plus de 4 bulletins de mariage même si seulement 4 sont décrits.
- *Pour les mariages légitimant un individu « EDP »* : utilisation des variables STATML et NBLEG.
- *Pour les naissances d'enfants d'individus « EDP »* : utilisation des variables STATNA(i) et OBNA(i).
- *Pour les décès d'individus « EDP »* : on ne prend en compte ici qu'un bulletin de décès par personne (144 personnes ont deux bulletins et une personne a 3 bulletins) ; utilisation des variables STATDC et OBDC.
- *Pour les reconnaissances* : utilisation des variables STATR(i) et PEMER(i) ; pour 3 716 individus, l'EDP compte plus de 2 bulletins de reconnaissance même si seulement 2 sont décrits.
- *Pour les bulletins d'enfant sans vie* : utilisation des variables STATMN(i) ; pour 2 individus, l'EDP compte plus de 2 bulletins d'enfants sans vie même si seulement 2 sont décrits.
- *Pour les bulletins d'enfant sans vie* : utilisation des variables TM(i) ; pour 18 individus, l'EDP compte plus de 2 mentions en marge même si seulement 2 sont décrites.

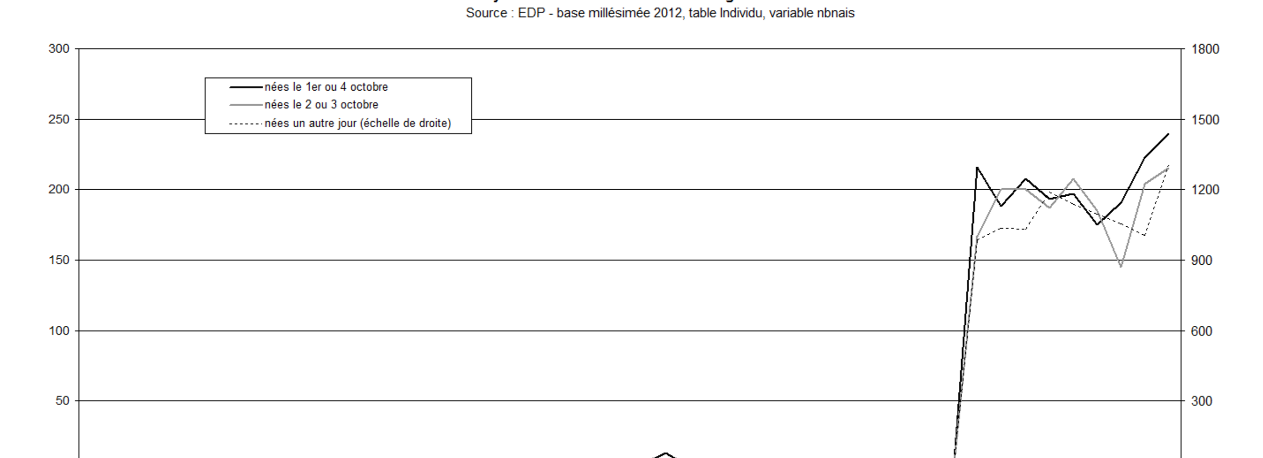
ANNEXE 3 - Quelques statistiques exploratoires

Cette annexe propose quelques statistiques exploratoires. Celles-ci sont notamment destinées à illustrer les informations démographiques que l'EDP permet de fournir ; à illustrer les effets des « trous EDP » sur ces informations afin de souligner l'importance de se limiter aux personnes nées les 1er et 4 octobre pour certaines analyses ; à donner un ordre de grandeur des effectifs disponibles selon différents critères.

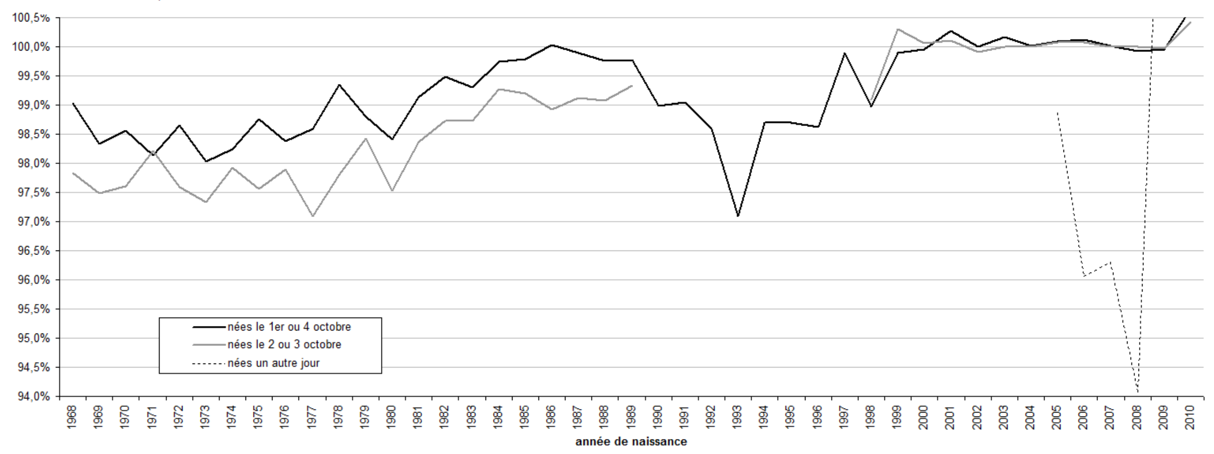
Les bulletins de naissance des personnes nées un jour « EDP »



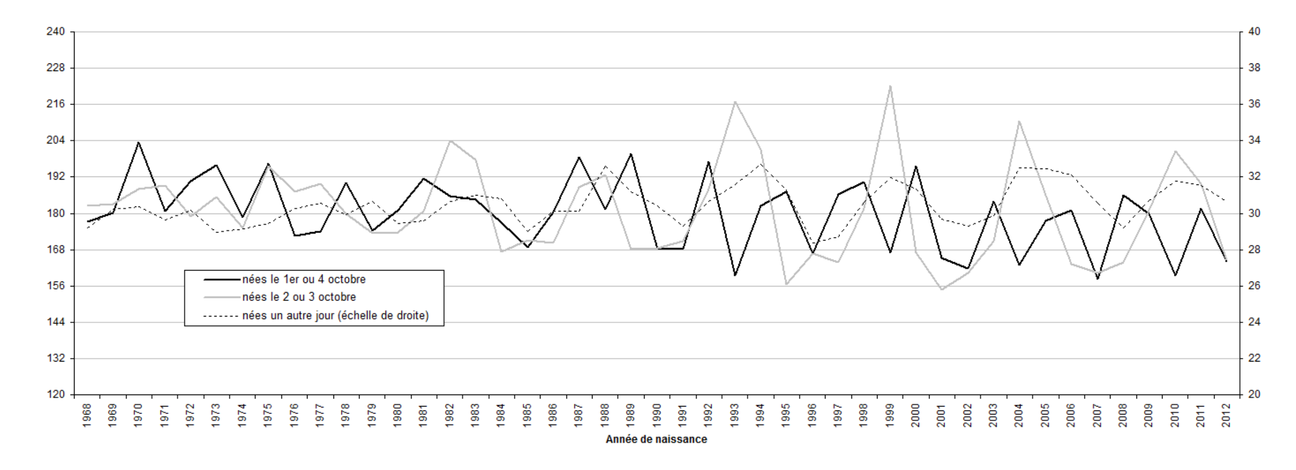
Note : les « nées un autre jour » sont les personnes nées les 12 nouveaux jours « EDP » introduits à partir de 2004 pour les évènements d'état-civil



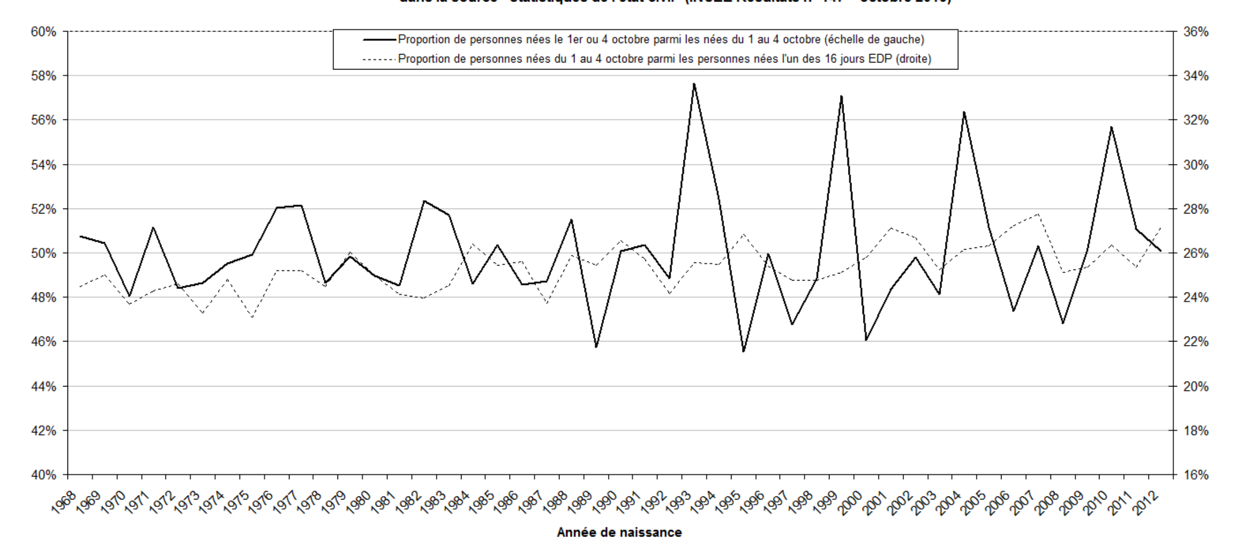
Note : les « nées un autre jour » sont les personnes nées les 12 nouveaux jours « EDP » introduits à partir de 2004 pour les évènements d'état-civil



Note : les « nées un autre jour » sont les personnes nées les 12 nouveaux jours « EDP » introduits à partir de 2004 pour les évènements d'état-civil

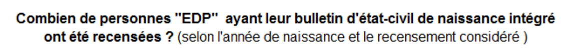


Note : les « nées un autre jour » sont les personnes nées les 12 nouveaux jours « EDP » introduits à partir de 2004 pour les évènements d'état-civil



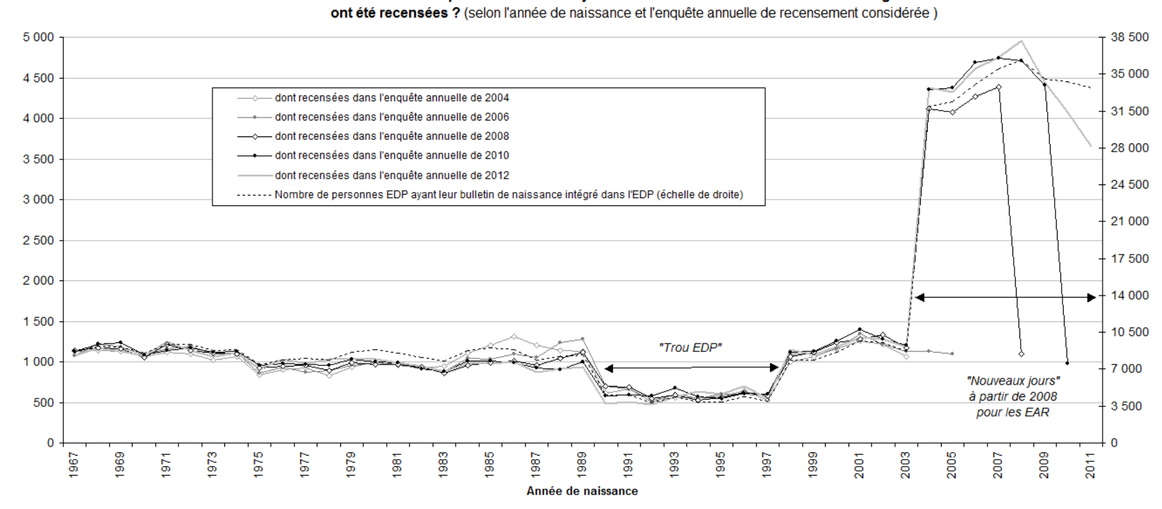
Note : les pics importants de fin de période pour la courbe noire correspondent aux années où le 1er octobre tombe un vendredi.

Combien de personnes « EDP » ayant leur bulletin de naissance intégrés ont aussi été recensés ?





Note : variables nai_date, nbnais, nbbi_rp68, nbbi_rp75, nbbi_rp82, nbbi_rp90, nbbi_rp99 (table Individu)
Source : Insee, Echantillon démographique permanent (base d'études 2012)
Champs : Nés en France métropolitaine (variable nai_lieu, table Individu)

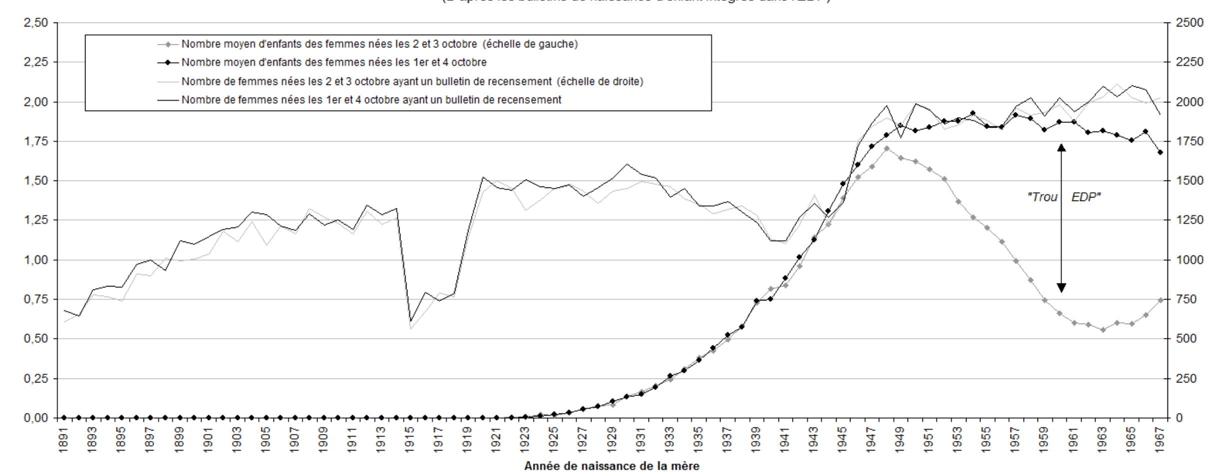


Note : variables nai_date, nbnais, nbbi_ear2004, nbbi_ear2006, nbbi_ear2008, nbbi_ear2010 (table Individu)
Source : Insee, Echantillon démographique permanent (base d'études 2012)
Champs : Nés en France métropolitaine (variable nai_lieu, table Individu)

La fécondité selon l'EDP

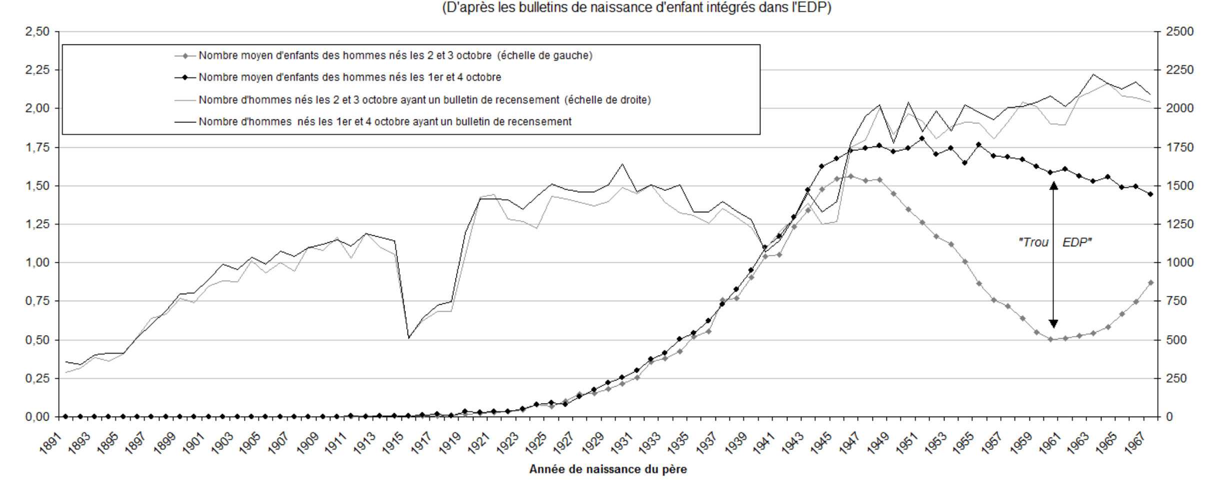
Fécondité des femmes et des hommes nés un jour EDP et recensés en 1968

Combien les femmes EDP recensées en 1968 ont eu d'enfants nés entre 1968 et 2012 ?
(D'après les bulletins de naissance d'enfant intégrés dans l'EDP)



Note : variables nai_date, nbenf, nbbi_rp1968, sexe (table Individu)
Source : Insee, Echantillon démographique permanent (base d'études 2012)
Champs : Nés en France métropolitaine (variable nai_lieu, table Individu)

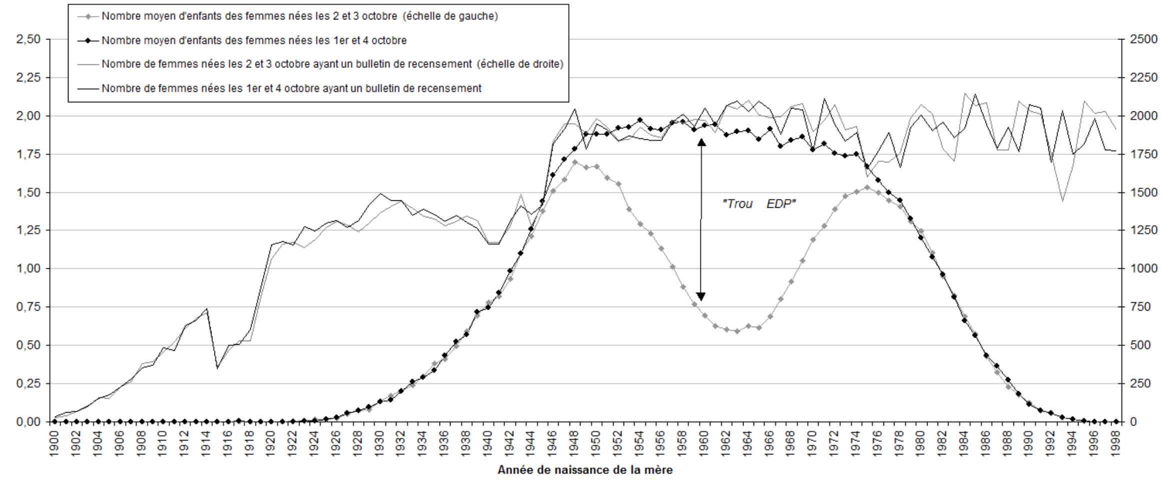
Combien les hommes EDP recensés en 1968 ont eu d'enfants nés entre 1968 et 2012 ?
(D'après les bulletins de naissance d'enfant intégrés dans l'EDP)



Note : variables nai_date, nbenf, nbbi_rp1968, sexe (table Individu)
Source : Insee, Echantillon démographique permanent (base d'études 2010)
Champs : Nés en France métropolitaine (variable nai_lieu, table Individu)

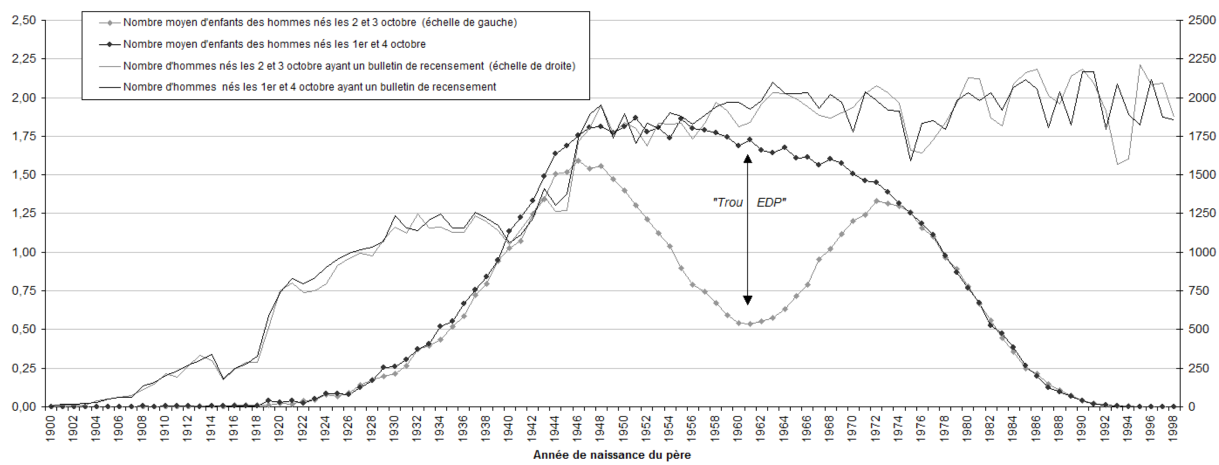
Fécondité des femmes et des hommes nés un jour EDP et recensés en 1999

Combien les femmes EDP recensées en 1999 ont eu d'enfants nés entre 1968 et 2012 ?
(D'après les bulletins de naissance d'enfant intégrés dans l'EDP)



Note : variables nai_date, nbenf, nbbi_rp1999, sexe (table Individu)
Source : Insee, Echantillon démographique permanent (base d'études 2012)
Champs : Nés en France métropolitaine (variable nai_lieu, table Individu)

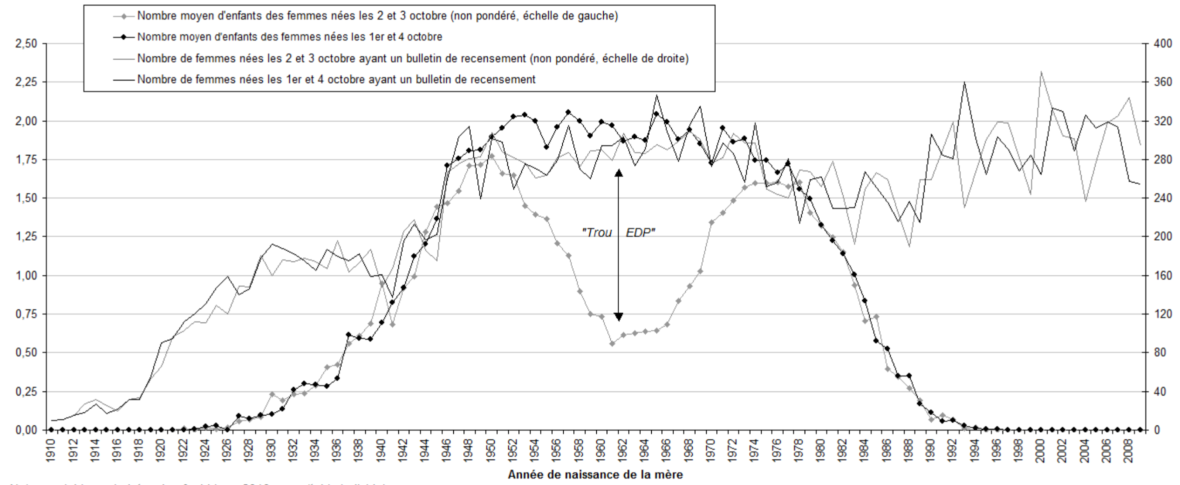
Combien les hommes EDP recensés en 1999 ont eu d'enfants nés entre 1968 et 2012 ?
(D'après les bulletins de naissance d'enfant intégrés dans l'EDP)



Note : variables nai_date, nbenf, nbbi_rp1999, sexe (table Individu)
Source : Insee, Echantillon démographique permanent (base d'études 2012)
Champs : Nés en France métropolitaine (variable nai_lieu, table Individu)

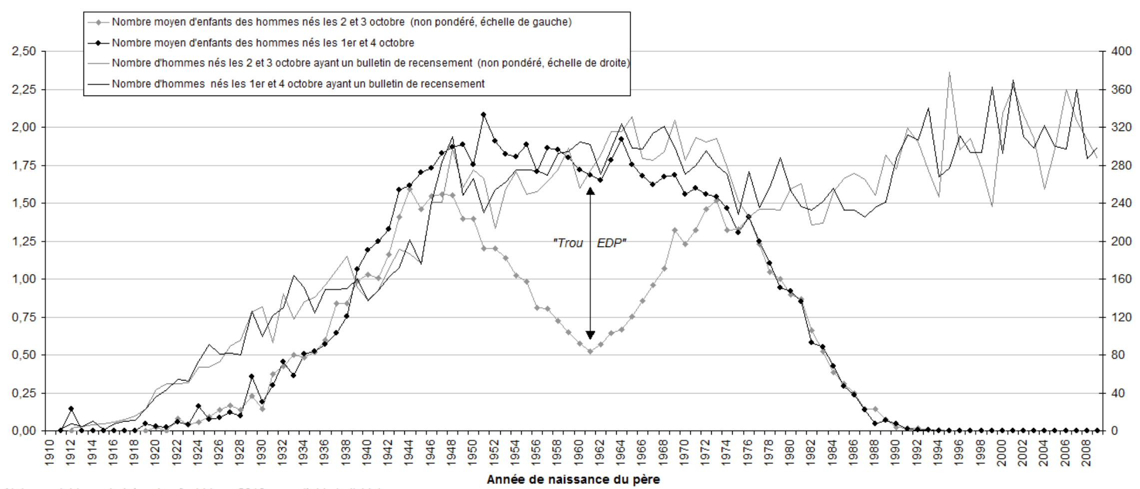
Fécondité des femmes et des hommes nés un jour EDP et recensés en 2010

Combien les femmes EDP recensées en 2010 ont eu d'enfants nés entre 1968 et 2012 ?
(D'après les bulletins de naissance d'enfant intégrés dans l'EDP)



Note : variables nai_date, nbenf, nbbi_ear2010, sexe (table Individu)
Source : Insee, Echantillon démographique permanent (base d'études 2012)
Champs : Nés en France métropolitaine (variable nai_lieu, table Individu)

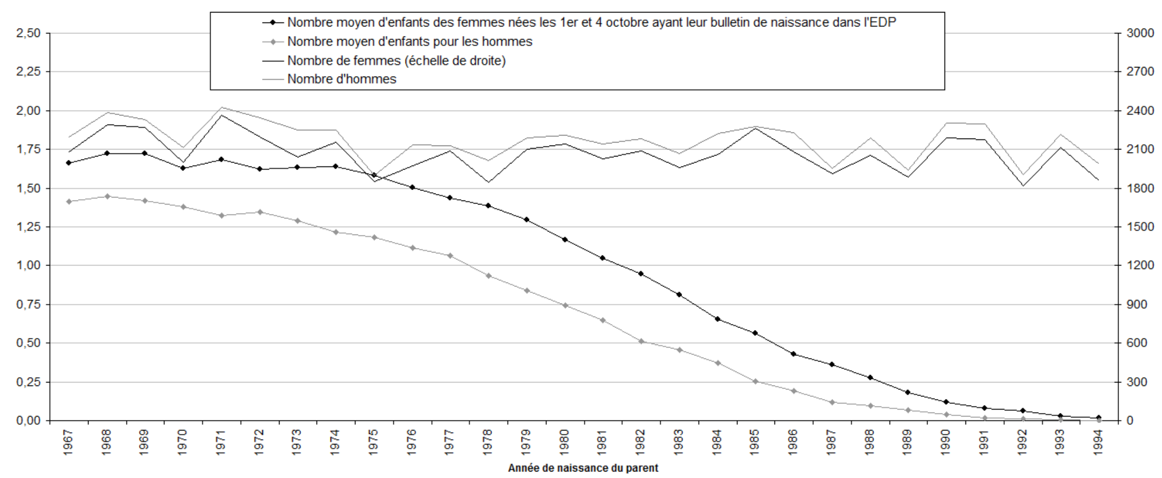
Combien les hommes EDP recensés en 2010 ont eu d'enfants nés entre 1968 et 2012 ?
(D'après les bulletins de naissance d'enfant intégrés dans l'EDP)



Note : variables nai_date, nbenf, nbbi_ear2010, sexe (table Individu)
Source : Insee, Echantillon démographique permanent (base d'études 2012)
Champs : Nés en France métropolitaine (variable nai_lieu, table Individu)

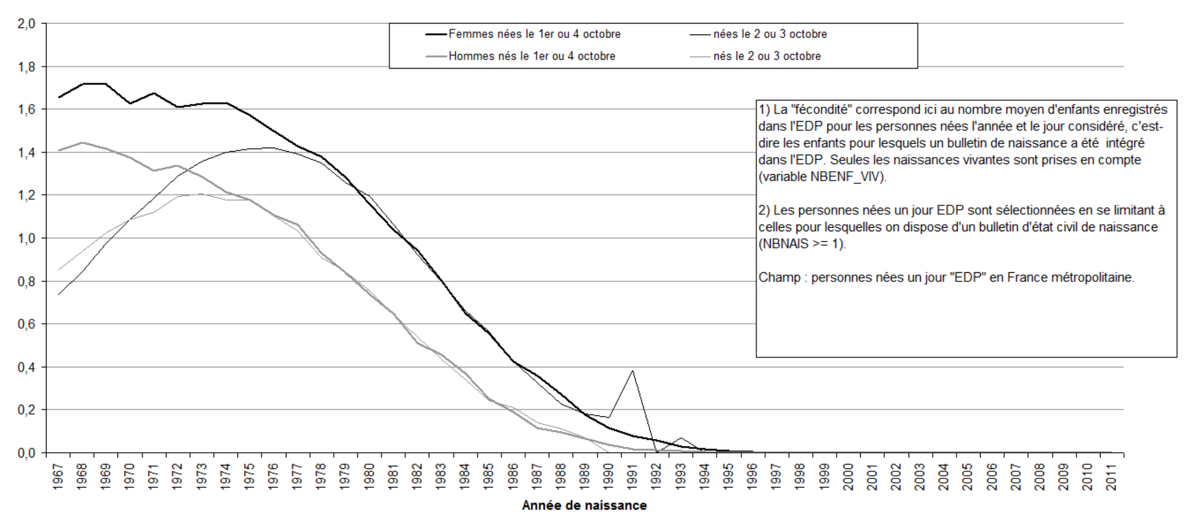
Nombre moyen d'enfants des personnes nées un jour EDP, selon leur année de naissance

Combien les personnes dont le bulletin de naissance est dans l'EDP ont eu d'enfants nés entre 1968 et 2012 ? (D'après les bulletins de naissance d'enfant intégrés dans l'EDP)

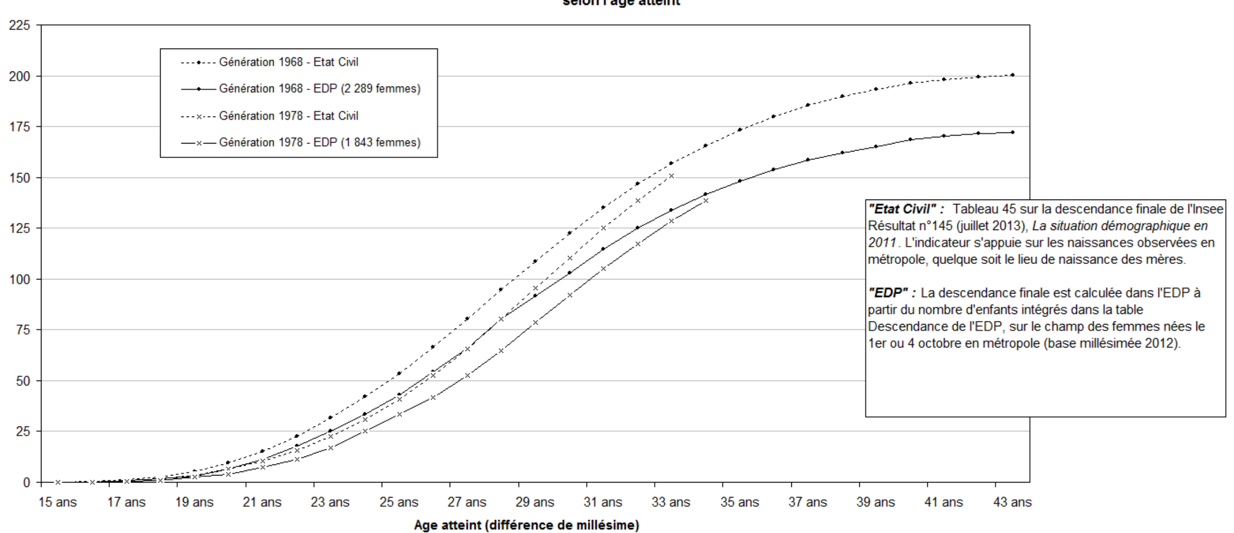


Note : Variables nai_date, nbenf, nbnais, sexe (table Individu)
Source : Insee, Echantillon démographique permanent (base d'études 2012)
Champs : Nés en France métropolitaine (variable nai_lieu, table Individu)

"Fécondité" observée des personnes "EDP" nées en France métropolitaine, selon leur date de naissance
Source : EDP - base millésimée 2012, table Individu.



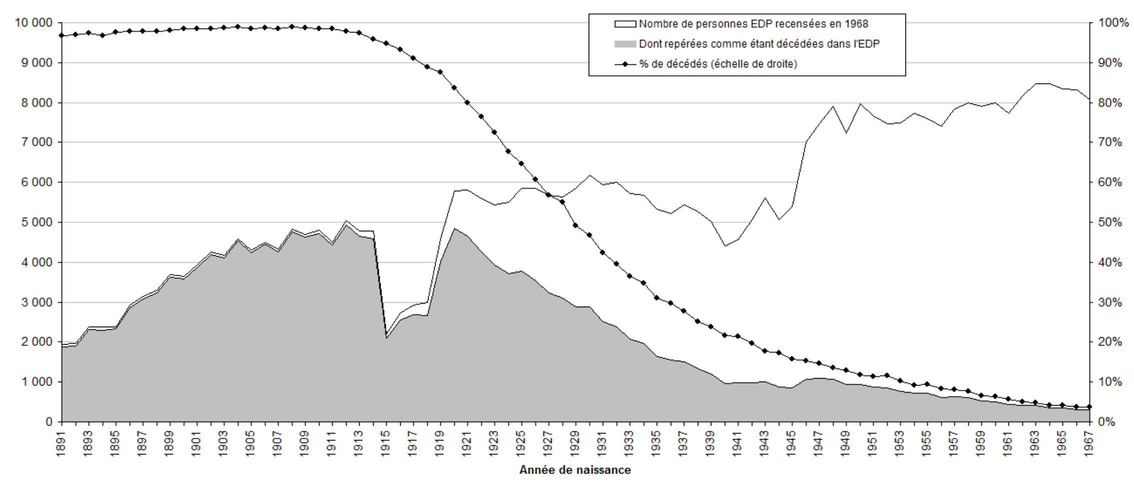
Comparaison de la descendance finale enregistrée dans l'EDP et de la descendance finale estimée à partir de l'état-civil



La mortalité selon l'EDP

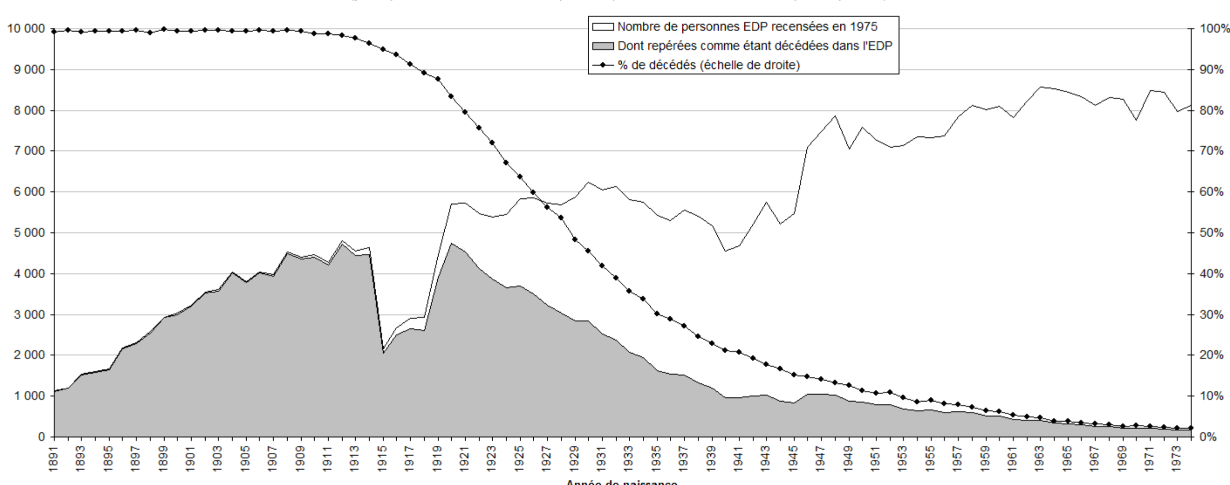
Pour différents recensements, proportion de personnes EDP repérées comme étant décédées entre le recensement et 2012

Combien de personnes EDP recensées en 1968 sont repérées comme étant décédées avant fin 2012 ?
(y compris les bulletins "fantômes" pour lesquels le bulletin de décès n'est pas disponible)



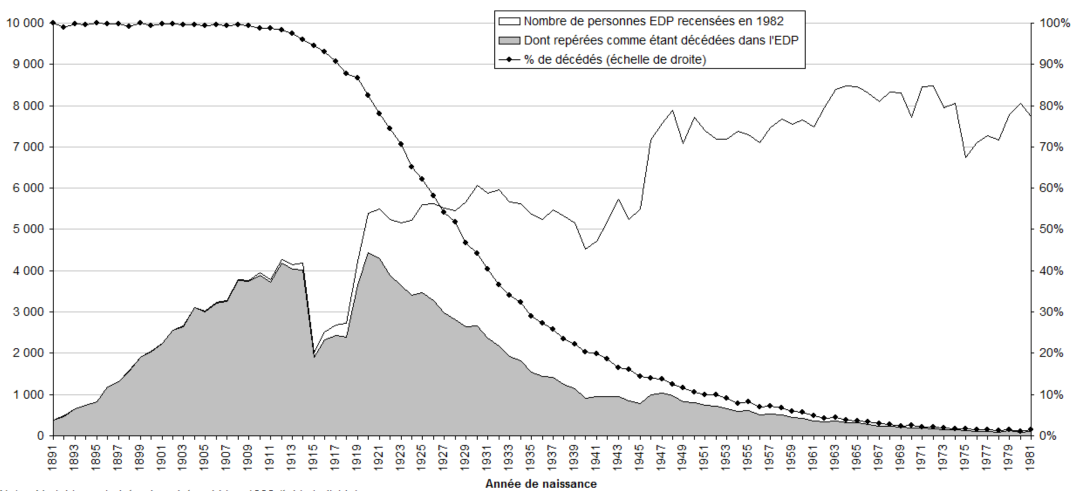
Note : Variables nai_date, dec_date, nbbi_rp1968 (table Individu)
Source : Insee, Echantillon démographique permanent (base d'études 2012)
Champs : Nés en France métropolitaine (variable nai_lieu, table Individu)

Combien de personnes EDP recensées en 1975 sont repérées comme étant décédées avant fin 2012 ?
(y compris les bulletins "fantômes" pour lesquels le bulletin de décès n'est pas disponible)



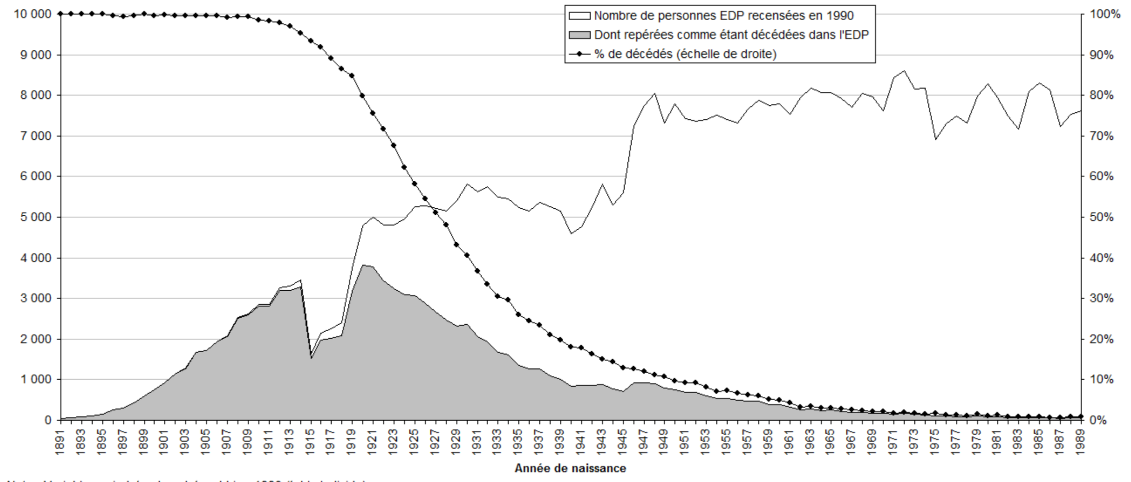
Note : Variables nai_date, dec_date, nbbi_rp1975 (table Individu)
Source : Insee, Echantillon démographique permanent (base d'études 2012)
Champs : Nés en France métropolitaine (variable nai_lieu, table Individu)

Combien de personnes EDP recensées en 1982 sont repérées comme étant décédées avant fin 2012 ?
(y compris les bulletins "fantômes" pour lesquels le bulletin de décès n'est pas disponible)



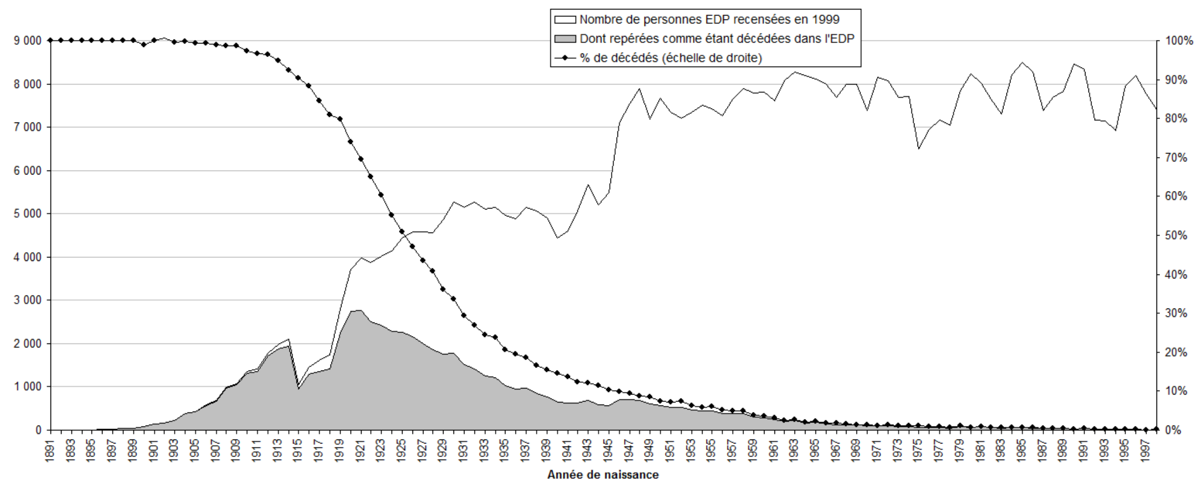
Note : Variables nai_date, dec_date, nbbi_rp1982 (table Individu)
Source : Insee, Echantillon démographique permanent (base d'études 2012)
Champs : Nés en France métropolitaine (variable nai_lieu, table Individu)

Combien de personnes EDP recensées en 1990 sont repérées comme étant décédées avant fin 2012 ?
(y compris les bulletins "fantômes" pour lesquels le bulletin de décès n'est pas disponible)



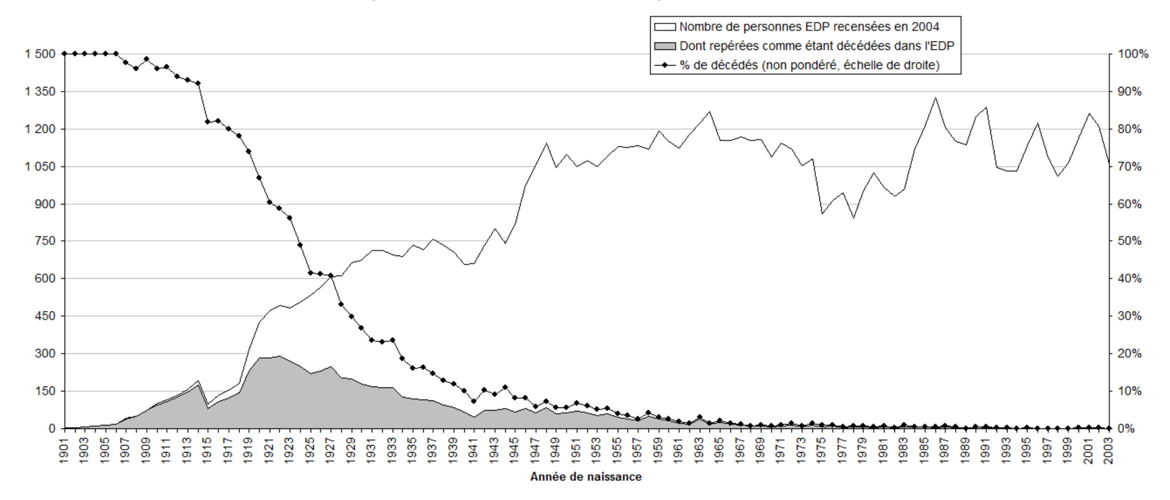
Note : Variables nai_date, dec_date, nbbi_rp1990 (table Individu)
Source : Insee, Echantillon démographique permanent (base d'études 2012)
Champs : Nés en France métropolitaine (variable nai_lieu, table Individu)

Combien de personnes EDP recensées en 1999 sont repérées comme étant décédées avant fin 2012 ?
(y compris les bulletins "fantômes" pour lesquels le bulletin de décès n'est pas disponible)



Note : Variables nai_date, dec_date, nbbi_rp1999 (table Individu)
Source : Insee, Echantillon démographique permanent (base d'études 2012)
Champs : Nés en France métropolitaine (variable nai_lieu, table Individu)

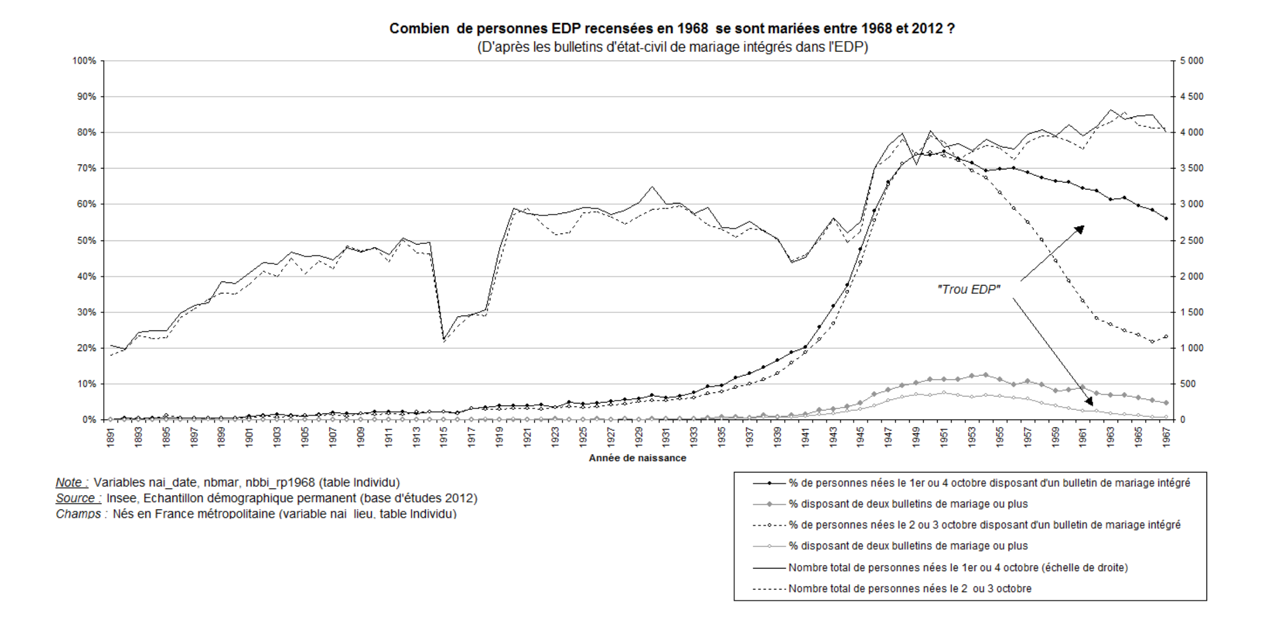
Combien de personnes EDP recensées en 2004 sont repérées comme étant décédées avant fin 2012 ?

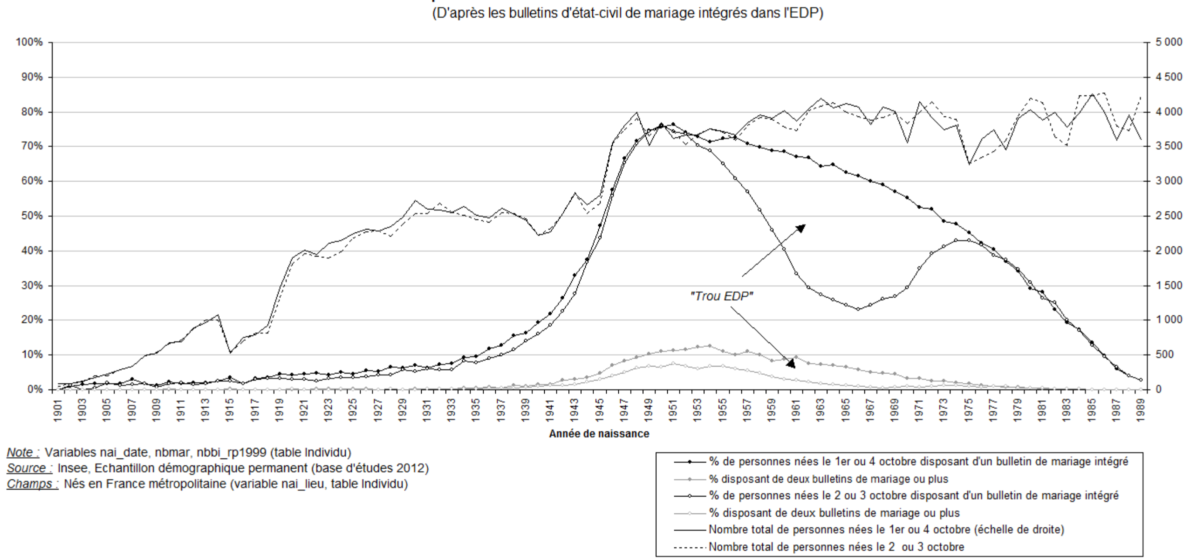


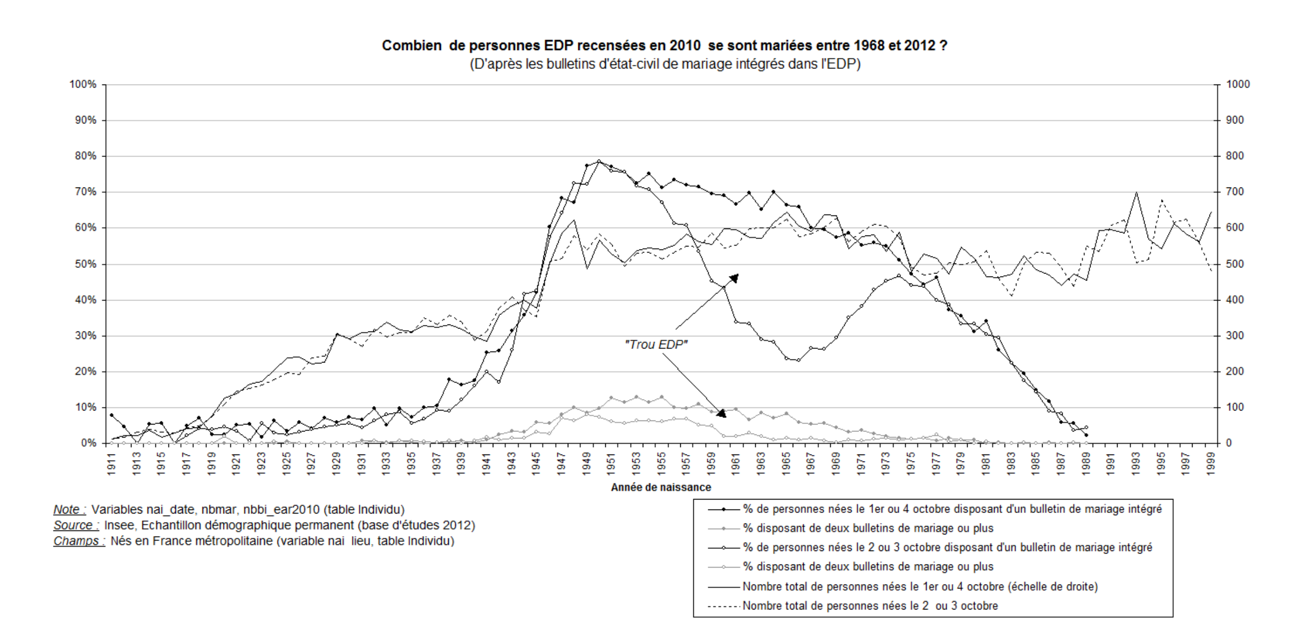
Note : Variables nai_date, dec_date, nbbi_ear2004 (table Individu)
Source : Insee, Echantillon démographique permanent (base d'études 2012)
Champs : Nés en France métropolitaine (variable nai_lieu, table Individu)

La nuptialité selon l'EDP

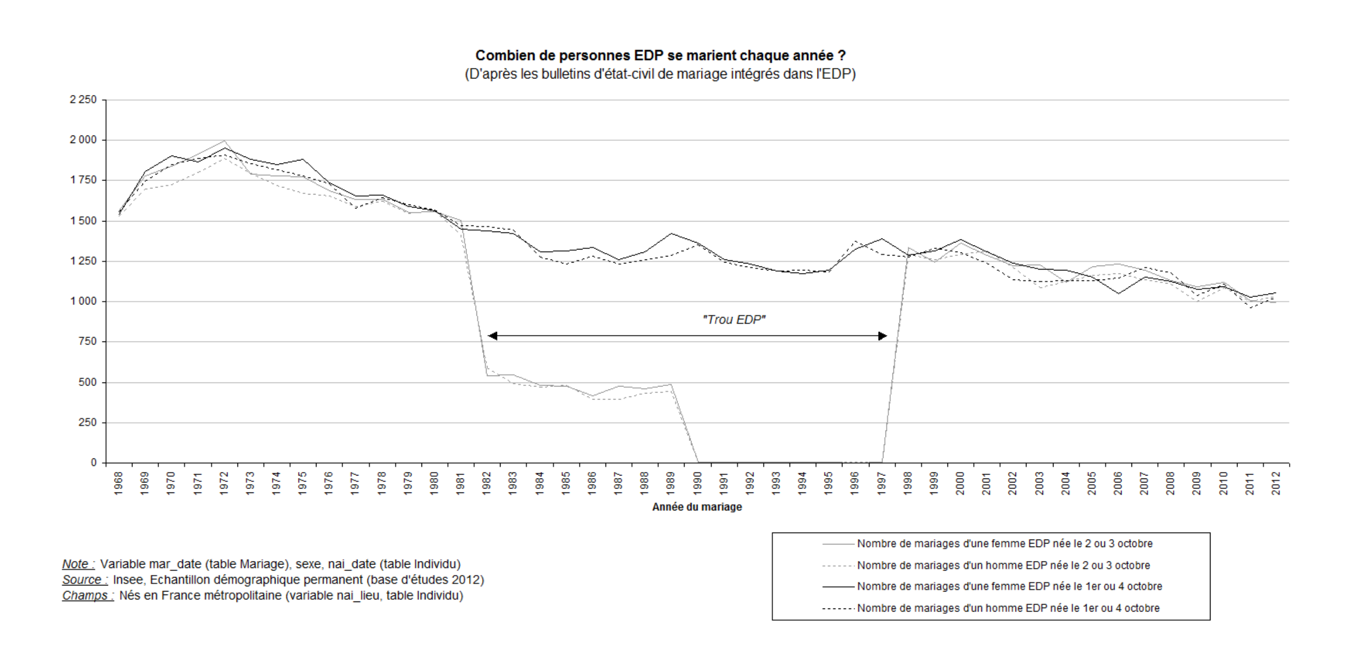
Proportion de personnes EDP recensées l'année considérée
pour lesquelles l'EDP dispose d'au moins un bulletin de mariage.



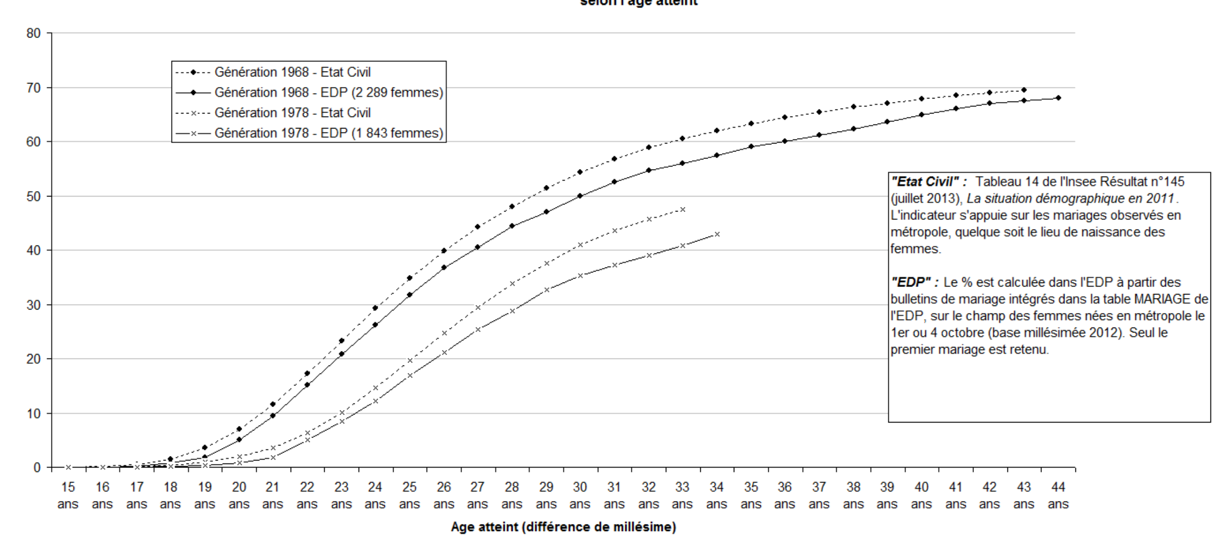




Chronique des bulletins de mariage disponibles dans l'EDP

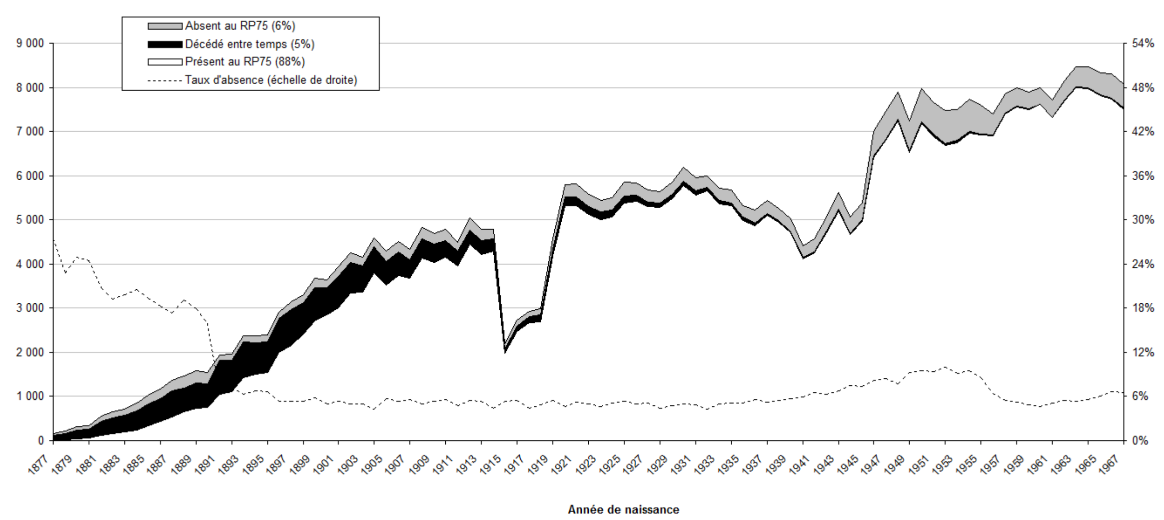


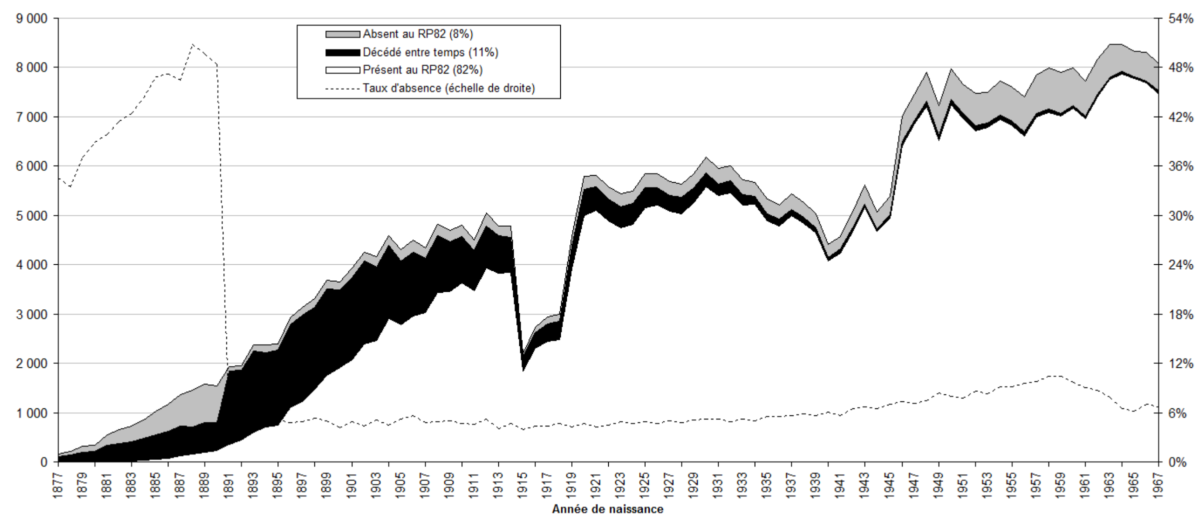
Comparaison de la nuptialité enregistrée dans l'EDP et de la nuptialité estimée à partir de l'état-civil



La présence d'un recensement à l'autre

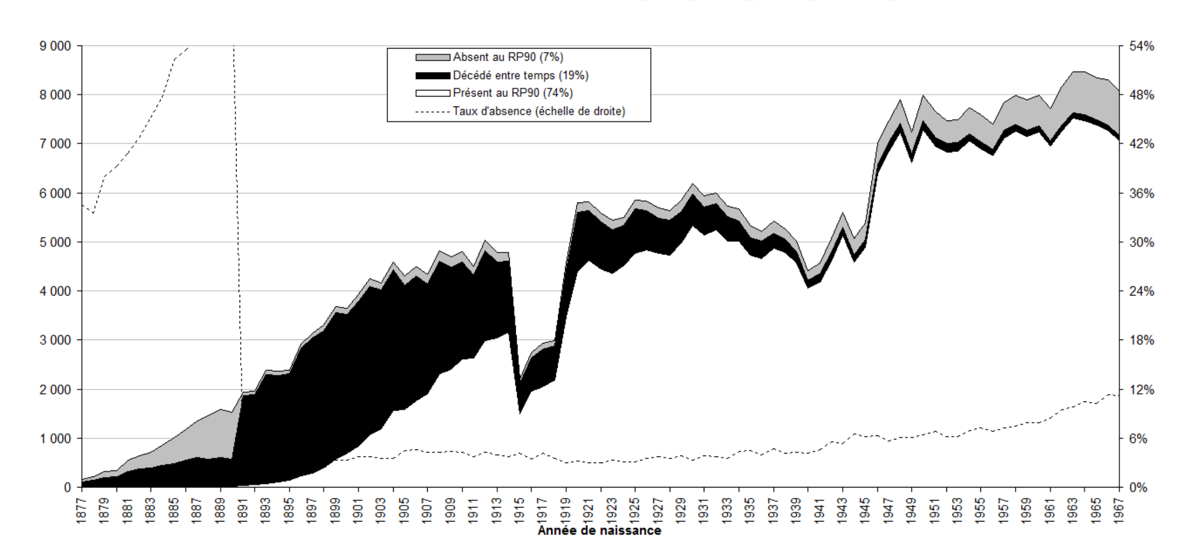
Combien de personnes EDP nées en métropole recensées en 1968 l'ont aussi été lors d'un recensement postérieur ?





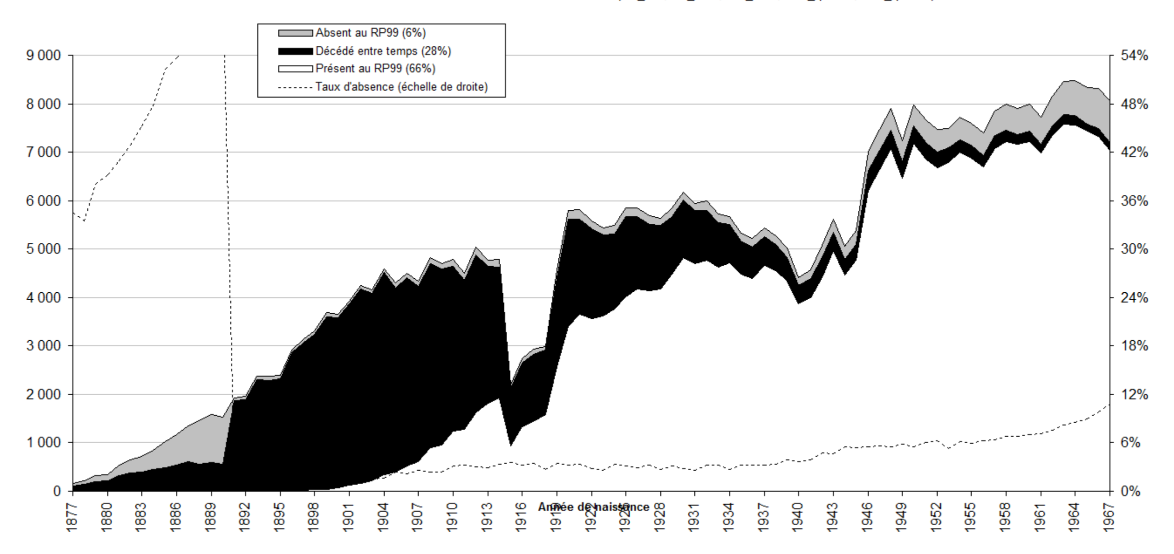
Les individus nés en métropole recensés en 1968 au recensement de 1990

Base d'études 2012. Variables de la table Individu (nai_lieu, nai_date, dec_date, nbbi_rp1968, nbbi_rp1990)

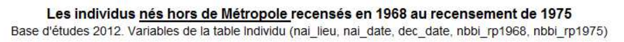


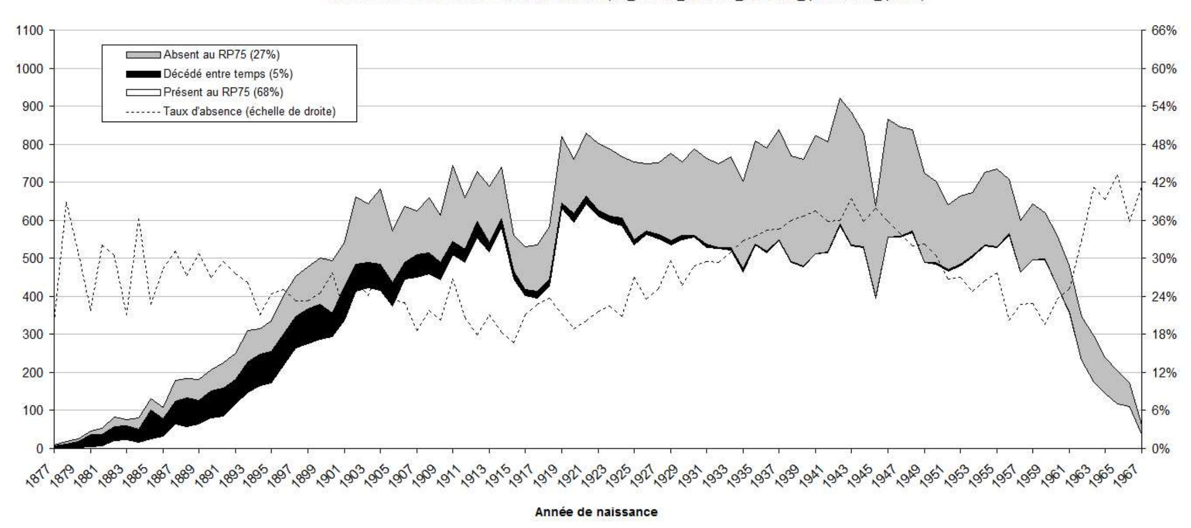
Les individus nés en métropole recensés en 1968 au recensement de 1999

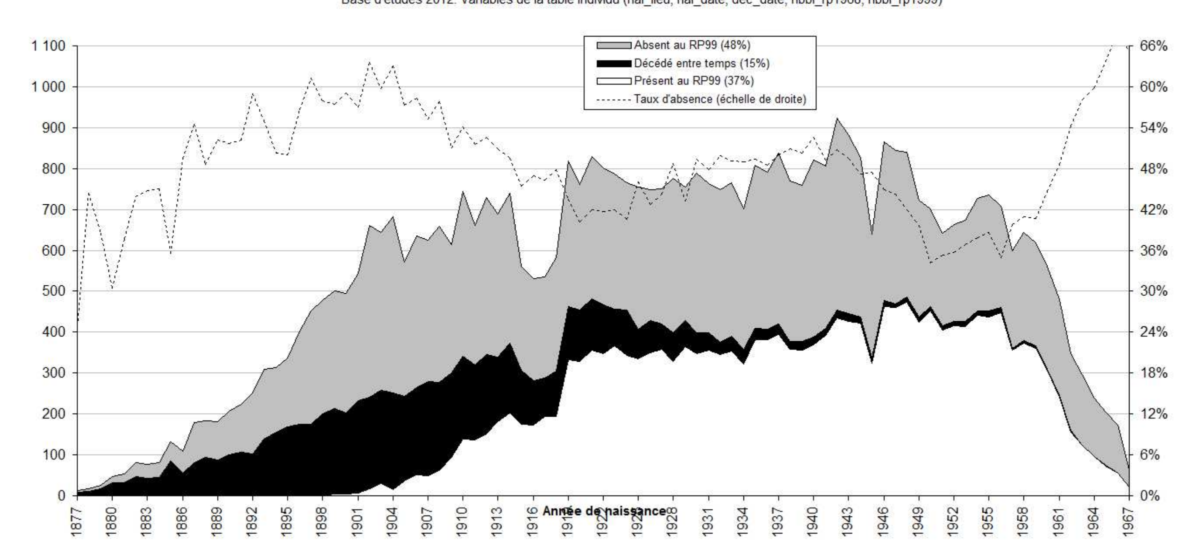
Base d'études 2012. Variables de la table Individu (nai_lieu, nai_date, dec_date, nbbi_rp1968, nbbi_rp1999)



Combien de personnes EDP <u>nées hors de métropole</u> recensées en 1968 l'ont aussi été lors d'un recensement postérieur ?







Combien de personnes EDP recensées en 1990 l'ont aussi été en 1999 ?

